JOURNAL OF
SYMPLECTIC GEOMETRY
Volume 13, Number 1, 151–207, 2015

Fukaya-Seidel category and gauge theory

Andriy Haydys

A new construction of the Fukaya-Seidel category associated with a symplectic Lefschetz fibration is outlined. Applying this construction in an infinite dimensional case, a Fukaya-Seidel-type category is associated with a smooth three-manifold. In this case the construction is based on a five-dimensional gauge theory.

1. Introduction

This paper consists of two major parts. In the first part, based on the idea of Seidel [24] we outline a construction of the Fukaya-Seidel category, which is associated with a symplectic manifold M equipped with the structure of a symplectic Lefschetz fibration. By this we mean, roughly speaking, a choice of an almost complex structure J and a J-holomorphic Morse function f. This construction does not rely on the notion of vanishing cycle but emphasizes instead the role of the antigradient flow lines of $\operatorname{Re}(e^{i\theta} f)$. In the second part, this construction is applied in the infinite dimensional case of the complex Chern-Simons functional. The corresponding construction conjecturally associates a Fukaya-Seidel-type category to a smooth three-manifold.

Our motivation originated from the suggestion to use higher dimensional gauge theory in studies of low dimensional manifolds as outlined in [14]. Namely, suppose we are given a construction that associates a higher dimensional manifold W_X to each lower dimensional manifold X from a suitable subclass and possibly equipped with an additional structure. The manifold W_X is assumed to be of dimension $6, 7$ or 8 and endowed with an $SU(3), G_2$ or $Spin(7)$ structure, respectively. Then, by counting higher dimensional instantons on W_X we should obtain an invariant of X. The construction studied in [14] in detail associates to each smooth spin four-manifold the total space of its spinor bundle.

Another construction of a similar nature associates to X^4 the total space of the "twisted spinor bundle" $\underline{\mathbb{R}} \oplus \Lambda^2_+ T^*X$. Then $Spin(7)$-instantons invariant along each fibre are solutions of the Vafa-Witten equations [27], while

$Spin(7)$-instantons invariant only along the fibres of $\Lambda^2_+ T^*X$ can be interpreted as antigradient flow lines of a function, whose critical points are solutions of the Vafa-Witten equations. It turns out that these flow lines can be obtained from certain elliptic equations on a general five-manifold W^5 equipped with a nonvanishing vector field by specializing to the case $W = X^4 \times \mathbb{R}$ just like flow lines of the real Chern-Simons functional are obtained from the anti-self-duality equations on $X^4 = Y^3 \times \mathbb{R}$. Specializing further to $W = Y^3 \times \mathbb{R}^2$ we obtain a construction of a Fukaya-type A_∞-category (this requires some extra choices) just like specialization of the anti-self-duality equations to $\Sigma^2 \times \mathbb{R}^2$ leads to the construction of the Fukaya A_∞-category associated with Σ. At this point an important distinction from the case of Riemann surfaces emerges. Namely, the construction involves a natural holomorphic function, the complex Chern-Simons functional, and this has significant implications for the flavour of the construction.

Having said this though, we do not appeal in this paper to higher dimensional anti-self-duality equations but rather begin directly with the formulation of the five-dimensional gauge theory. From this perspective the most interesting theory is obtained via reduction to three-manifolds, where the construction of the A_∞-category admits a finite-dimensional interpretation in the framework of symplectic geometry.

This paper is organized as follows. In Section 2 we describe the construction of the Fukaya-Seidel category in the finite-dimensional case. From one point of view, this construction is a generalization of a Floer theory, where *generators* of the homology groups are antigradient flow lines of the real part of a holomorphic Morse function connecting a pair of critical points. Then the Floer differential is obtained from pseudoholomorphic planes with a Hamiltonian perturbation satisfying certain asymptotic conditions (see (2.7)– (2.9) for more details).

Sections 3 and 4 are devoted to the formulation of the five-dimensional gauge theory and its various dimensional reductions. In Section 5 we describe applications of the equations obtained in the previous sections to low dimensional topology. In particular, one can (conjecturally) associate an integer to a five-manifold, Floer-type homology groups to a four-manifold and a Fukaya-Seidel-type category to a three-manifold. In dimension three, critical points correspond to flat G^c-connections on Y, flow lines correspond to Vafa-Witten-type instantons on $Y \times \mathbb{R}$ and pseudoholomorphic planes correspond to "five-dimensional instantons" on $Y \times \mathbb{R}^2$. This should be a part of a multi-tier (extended) quantum field theory [12] but we do not study this aspect in the current paper.

The constructions described in this paper may also be useful in other settings, for instance in the context of Calabi-Yau threefolds. Here the critical points of the holomorphic Chern-Simons functional correspond to holomorphic vector bundles over a Calabi-Yau threefold Z, flow lines correspond to G_2-instantons on $Z \times \mathbb{R}$ and pseudoholomorphic planes correspond to $Spin(7)$-instantons on $Z \times \mathbb{R}^2$.

Many aspects of this paper are related to ideas of various authors. As it has been already mentioned above, our construction of the Fukaya-Seidel category in the finite dimensional case is a modification of Seidel's idea. The equation we utilize for the definition of the structure maps in the Fukaya-Seidel A_∞-category was used in the context of mirror symmetry in [9] ("Witten equation") in the case of quasi-homogeneous polynomials. The antigradient flow lines of the real part of the holomorphic Chern-Simons functional appeared in [16] for the first time and were further studied in [28, 29]. In [8] antigradient flow lines of the real part of the holomorphic Chern-Simons functional are used in the context of Calabi-Yau threefolds.

2. Fukaya-Seidel categories of symplectic Lefschetz fibrations

In [24, 25] Seidel describes the construction of a Fukaya category associated with a symplectic Lefschetz fibration in terms of vanishing cycles. In the first part of this section we describe omitting (important) technical details an alternative approach, which does not rely on the notion of vanishing cycle. The rest of the section is devoted to basic analytic properties of the objects involved in the construction.

2.1. Symplectic Lefschetz fibrations

Let $(M^{2n}, \omega, \lambda)$, $\omega = d\lambda$, be an exact symplectic manifold with boundary. Choose an almost complex structure J such that $g(\cdot,\cdot) = \omega(\cdot, J\cdot)$ is a Riemannian metric on M. It is also convenient to assume that J is orthogonal with respect to g. Let $f\colon M \to \mathbb{C}$ be a J-holomorphic function. We assume the following properties:

(P1) f is a proper map with finitely many non-degenerate critical points lying in pairwise different fibres. Moreover, locally near each critical point J is integrable.

(P2) The boundary of M is convex [17, Def. 9.2.5].

(P3) Let $M_0 = f^{-1}(z_0)$ be a regular fiber. Then there exist compact subsets $K \subset M \setminus \partial M$, $K' \subset M_0 \times \mathbb{C} \setminus \partial M_0 \times \mathbb{C}$, and a positive number r with the following significance. Denote $V = M \setminus K$, $V' = M_0 \times \mathbb{C} \setminus K'$. Then for each $z \in \mathbb{C}$ there exists a small neighbourhood $B_\delta(z)$ and a fiber preserving diffeomorphism ψ_z such that the following holds: The diagram

$$\begin{array}{ccc} V \cap \big(M_0 \times B_\delta(z)\big) & \xrightarrow{\psi_z} & V' \cap f^{-1}\big(B_\delta(z)\big) \\ & {\scriptstyle pr_2} \searrow \quad \swarrow {\scriptstyle f} & \\ & B_\delta(z) & \end{array}$$

commutes, ψ_z is the identity on $(M_0 \times \{z_0\}) \cap V$ whenever $z_0 \in B_\delta(z)$, and the pull-back of (λ, J) is $(\lambda_{M_0} + r\lambda_0, J_{M_z} \times I_0)$. Here $\lambda_0 = \mathrm{Re}\,(iz d\bar{z})$ is the primitive of the standard symplectic form ω_0 and I_0 is the standard complex structure on $\mathbb{C}$.

It is worth pointing out that properties (P1) and (P3) imply that there exists $R > 0$ such that the preimage of $B_R^c(0)$ is contained in V, where $B_R^c(0)$ denotes the complement of the ball $B_R(0)$ of radius R. In other words, for any $z \in B_R^c(0)$ there exists a neighbourhood $B_\delta(z) \subset B_R^c(0)$, and fiber preserving diffeomorphism $\psi_z \colon M_0 \times B_\delta(z) \to f^{-1}(B_\delta(z))$ with the properties as in (P3). Similarly, there exists a neighbourhood W of ∂M, a neighbourhood W' of $\partial(M_0 \times \mathbb{C}) = \partial M_0 \times \mathbb{C}$, and a diffeomorphism $\psi \colon W' \to W$ such that the pull-back of (λ, J) is $(\lambda_{M_0} + r\lambda_0, J_{M_0} \times I_0)$. Here M_0 is some fiber. Conversely, these two properties imply (P3).

Denote

$$f = f_0 + i f_1, \qquad \rho = \{f_0, f_1\}.$$

Another consequence of (P3) is that $\rho = r^{-1}$ on V. In particular, this implies that ρ is bounded on M.

The following interpretation of ρ will be useful in the sequel. Denote

$$v_0 = \mathrm{grad}\, f_0 \quad \text{and} \quad v_1 = \mathrm{grad}\, f_1.$$

The holomorphicity of f implies $Jv_0 = v_1$. Then the Hamiltonian vector field of f_0 is $X_{f_0} = -Jv_0 = -v_1$. This yields

$$\rho(m) = |v_0(m)|^2 = |v_1(m)|^2 \qquad \text{for } m \in M. \tag{2.1}$$

Remark 2.1. It is interesting to notice that property (P3) is in fact equivalent to ρ being constant on a complement of a compact subset. Indeed, assume ρ is constant on $V = M \setminus K$. Then the identity $[v_0, v_1] = -[X_{f_1}, X_{f_0}] = X_{\{f_0,f_1\}}$ implies that v_0 and v_1 commute on V. The subset $M \setminus \mathrm{Crit}(f)$ is equipped with the connection, which is induced by the symplectic form. Then $\rho^{-1}v_0$ and $\rho^{-1}v_1$ are the horizontal lifts of $\frac{\partial}{\partial s}$ and $\frac{\partial}{\partial t}$, respectively, where (s,t) be coordinates on $\mathbb{C} \cong \mathbb{R}^2$. Hence, the connection is flat over V. It follows that in a flat trivialization in a neighbourhood of some $z \in \mathbb{C}$ the symplectic form can be written as $\omega_{M_z} + r\omega_0$, where r is some function. Then r is constant, since $r^{-1} = \rho$.

Examples of the fibrations with properties (P1)–(P3) can be found in [25, (19b)] (it is only needed to drop the restriction to the preimage of a large disc).

Other examples can be constructed starting from symplectic Lefschetz fibrations over the disc $\pi\colon E \to D = B_1(0)$ as in [25, (15a)] assuming triviality near the horizontal boundary [25, Remark 15.2]. Indeed, first of all on an open neighbourhood of $\partial^h E$ diffeomorphic to an open neighbourhood of the horizontal boundary of the trivial fibration $E_{pt} \times D$ with the help of a suitable cut-off function we can deform the symplectic form to $\omega_{E_{pt}}$. This is clearly no longer symplectic form on the horizontal subbundle but later on we will add some multiple of the standard symplectic form on D so that the resulting 2-form will be symplectic on the total space.

To extend E to a fibration over the whole complex plane proceed as follows. Choose $\delta > 0$ such that all critical values of π are contained in $B_{1-\delta}(0)$. With the help of the parallel transport along radial lines we obtain

$$E\big|_Z \cong pr^*E\big|_{S^1_{1-\delta}}, \tag{2.2}$$

where $pr\colon Z = \{1-\delta \le |z| \le 1\} \cong S^1_{1-\delta} \times [1-\delta, 1] \to [1-\delta, 1]$. Denote by (ϱ, φ) the polar coordinates and by $E_{1-\delta,0}$ the fiber over the point $(\varrho,\varphi) = (1-\delta, 0)$. Then we can write [25, (15a)] the symplectic 2-form on $pr^*E\big|_{S^1_{1-\delta}}$ in the form

$$\omega = \omega_{E_{1-\delta,0}} + d\kappa,$$

where $\kappa = \kappa_1(\varrho,\varphi)d\varrho + \kappa_2(\varrho,\varphi)d\varphi + dR$ for some functions $\kappa_1, \kappa_2, R \in C^\infty([1-\delta, 1] \times S^1 \times E_{1-\delta,0})$ (the notation does not reflect the dependence on all variables). Choose smooth cut-off functions $\alpha, \beta\colon [1-\delta, +\infty] \to [0,1]$ such that

$$\alpha(\varrho) = \begin{cases} \varrho & \varrho \in [1-\delta, 1-\frac{2\delta}{3}], \\ 1 & \varrho \ge 1-\frac{\delta}{3}, \end{cases} \qquad \beta(\varrho) = \begin{cases} 1 & \varrho \in [1-\delta, 1-\frac{2\delta}{3}], \\ 0 & \varrho \ge 1-\frac{\delta}{3}, \end{cases}$$

and denote $\kappa' = \kappa_1(\alpha(\varrho), \varphi)\, d\varrho + \kappa_2(\alpha(\varrho), \varphi)\, d\varphi + d(\beta R)$. This defines a connection 1-form on

$$E\big|_{S^1_{1-\delta}} \times [1-\delta, +\infty) \xrightarrow{\ p\ } \big\{|z| > 1-\delta\big\}. \tag{2.3}$$

Then for sufficiently large $r > 0$ the 2-form $\omega_r = \omega_{E_{1-\delta,0}} + d\kappa' + rp^*\omega_0$ is symplectic and equals to $\omega_E + r\pi^*\omega_0$ over $\{1-\delta < |z| < 1 - \frac{2\delta}{3}\}$. Hence, $E\big|_{B_{1-\delta}(0)}$ can be glued with (2.3) to obtain a fibration over the whole complex plane. By construction, this has properties (P1)–(P3).

2.2. Outline of the construction

The purpose of this subsection is to outline the main points of the alternative construction of the Fukaya-Seidel A_∞-category. The discussion of technical details is postponed to the proceeding subsections.

Let us briefly recall the basic ingredients of the Fukaya-Seidel A_∞-category (see [24, 25] for details). For the sake of simplicity we consider the ungraded version with coefficients in $\mathbb{Z}/2\mathbb{Z}$ ("preliminary version" in the terminology of [25]). It is convenient to choose a basepoint z_0, which does not lie on any straight line determined by a pair of critical values (in particular, z_0 is distinct from critical values). Denote by $m_1, \dots, m_k$ critical points of f and put $z_j = f(m_j)$. The indexing can be chosen such that the sequence $\arg(z_j - z_0) \in (-\pi, \pi]$ is decreasing in j and this defines a linear order on the set of critical points.

Choose a collection of paths connecting z_0 with each z_j missing the remaining critical values. Let $L_j \subset f^{-1}(z_0)$ be the vanishing cycle of m_j associated with the path connecting z_0 and z_j. Denote by Γ the ordered collection $(L_1, \dots, L_k)$. Seidel associates to Γ a directed Fukaya A_∞-category $Lag^{\to}(\Gamma)$, whose objects are vanishing cycles L_j and morphisms are Floer chain complexes as follows. First recall that an A_∞-structure is a collection of maps

$$\begin{aligned}\mu^d\colon hom(L_{j_d}, L_{j_{d+1}}) \otimes \cdots \otimes hom(L_{j_1}, L_{j_2}) &\longrightarrow hom(L_{j_1}, L_{j_{d+1}}),\\ d = 1, 2, 3, \dots\end{aligned}$$

satisfying certain quadratic relations [25, (1.2)] and by the directedness we have

$$hom(L_j, L_k) = \begin{cases} CF(L_j, L_k) & j < k,\\ \mathbb{Z}/2 \cdot id & j = k,\\ 0 & j > k.\end{cases}$$

The Floer complex $CF(L_j, L_k)$ is generated by the points of $L_j \cap L_k$ and the map μ^1 is the Floer differential, which counts pseudoholomorphic strips such that one boundary component is mapped to L_j and the other component is mapped to L_k. The maps μ^d for $d \geq 2$ are defined similarly by counting pseudoholomorphic discs with $d+1$ punctures on the boundary. The resulting A_∞-category $Lag^{\rightarrow}(\Gamma)$ depends on the choices made but Seidel shows that the derived category $D^b(Lag^{\rightarrow}(\Gamma))$ is an invariant of the Lefschetz fibration.

With this understood we now give another construction of the Fukaya-Seidel A_∞-category. Pick a pair of critical points (m_-, m_+) and denote $\theta_\pm = \arg(z_\pm - z_0) \in (-\pi, \pi]$. Let $\gamma_m^\pm$ be the solution of the Cauchy problem

$$\dot\gamma_m^\pm + \cos\theta_\pm\, v_0 + \sin\theta_\pm\, v_1 = 0, \quad \gamma_m^\pm(0) = m \in f^{-1}(z_0).$$

Notice that the image of $f \circ \gamma_m^\pm \colon \mathbb{R} \to \mathbb{C}$ is contained in a straight line passing through z_0 and $z_\pm$. Then the vanishing cycle $L_\pm$ of $m_\pm$ associated with the segment $\overline{z_0 z_\pm}$ can be conveniently described as

$$L_\pm = \big\{ m \in f^{-1}(z_0) \;\big|\; \lim_{t\to+\infty} \gamma_m^\pm(t) = m_\pm \big\}.$$

Then, if we denote

$$\theta_0(t) = \begin{cases} \theta_+ & t \leq 0, \\ \arg i(z_- - z_0) = \theta_- \pm \pi & t > 0, \end{cases} \tag{2.4}$$

the set $L_+ \cap L_-$ can be identified with the space of solutions of the problem

$$\dot\gamma + \cos\theta_0(t)\, v_0 + \sin\theta_0(t)\, v_1 = 0, \qquad \lim_{t\to\pm\infty} \gamma(t) = m_\mp. \tag{2.5}$$

Here solutions are understood to be smooth on $\mathbb{R} \setminus \{0\}$ and continuous at $t = 0$. We call solutions of (2.5) *broken flow lines* of f connecting m_- and m_+ and denote by $\Gamma_0(m_-; m_+)$ the space of all solutions. Notice that for each broken antigradient flow line γ the image of $f \circ \gamma$ lies on the curve $\overline{z_- z_0 z_+}$ and $f \circ \gamma(0) = z_0$.

It will be convenient in the sequel to replace θ_0 by a smooth function θ_ν, where ν is a real parameter. The choice of the function θ_ν, which is described in Subsection 2.3 in details, turns out to be quite important, but what we need to know at this point is that θ_ν is close to θ_0 for ν small enough.

Denote by $\Gamma_\nu = \Gamma_\nu(m_-, m_+)$ the space of solutions of the problem

$$\dot\gamma + \cos\theta_\nu(t)\, v_0 + \sin\theta_\nu(t)\, v_1 = 0, \qquad \lim_{t\to\pm\infty} \gamma(t) = m_\mp. \tag{2.6}$$

We also call solutions of Equations (2.6) broken flow lines.

Remark 2.2. We assume that for ν small enough there exists a correspondence between solutions of (2.6) and (2.5). This is discussed in detail in Appendix B. The advantage of Problem (2.6) is that its solutions are smooth everywhere on $\mathbb{R}$.

Furthermore, notice that the Floer differential μ^1 should take broken flow lines as input and should return formal linear combinations of broken flow lines as output. With $m_\pm$ as above, pick additionally two solutions $\gamma_\pm$ of Equations (2.6). Then the role of holomorphic strips with boundary on $L_\pm$ in our framework is played by solutions of the problem

$$\partial_s u + J\big(\partial_t u + \cos\theta_\nu\, v_0 + \sin\theta_\nu\, v_1\big) = 0, \qquad u\colon \mathbb{R}^2_{s,t} \to M, \tag{2.7}$$

$$\lim_{t\to\pm\infty} u(s,t) = m_\mp, \qquad \lim_{t\to\pm\infty} \int_{-\infty}^{+\infty} |\partial_s u(s,t)|\, ds = 0, \tag{2.8}$$

$$\lim_{s\to\pm\infty} u(s,t) = \gamma_\mp(t), \qquad \lim_{s\to\pm\infty} \int_a^b |\partial_s u(s,t)|\, dt = 0. \tag{2.9}$$

Here the limits appearing on the left hand side of (2.8) and (2.9) are understood in the $C^0(\mathbb{R})$-topology and $a \le b$ are arbitrary. Notice that (2.7) is the pseudoholomorphic map equation with a Hamiltonian perturbation. Namely, the time-dependent Hamiltonian function here is $\operatorname{Im}(e^{-i\theta_\nu(t)} f)$.

Notice also that it is assumed that the integral in (2.8) is convergent for all $t \in \mathbb{R}$. For instance, this is the case if $\partial_s u \in W^{k,p}(\mathbb{R}^2; u^*TM)$ with $k > \max\{\frac{1}{p}, \frac{2}{p} - 1\}$. In this case, for any fixed τ we have

$$\|\partial_s u(\cdot,\tau)\|_{L^1(\mathbb{R})} \le C_{k,p} \|\partial_s u\|_{W^{k,p}(H_\tau)},$$

where $H_\tau = \{t \ge \tau\} \subset \mathbb{R}^2$. In particular, $\int_{-\infty}^{+\infty} |\partial_s u(s,t)|\, ds$ tends to zero as $t \to +\infty$ and similarly for $t \to -\infty$.

It is very instructive to see a relation between solutions of (2.7)–(2.9) and pseudoholomorphic strips as in Seidel's approach. This is outlined in Appendix A. However, instead of proving that such a connection indeed holds, we study Equations (2.7)–(2.9) directly, since in view of the intended applications it is important to have direct proofs of the basic properties (compactness, Fredholm property, transversality etc.). In this paper we prove compactness and Fredholm property for solutions of (2.7)–(2.9).

Next we show how to define the map μ^2 in our framework. Let Ω be a (non-compact) Riemann surface containing three "long necks". By this we mean a triple of holomorphic embeddings

$$\imath_1, \imath_2\colon \{z \mid \operatorname{Re} z < 0\} \to \Omega, \qquad \text{and} \qquad \imath_3\colon \{z \mid \operatorname{Re} z > 0\} \to \Omega$$

with disjoint images. To be more explicit, we choose the complex plane $\mathbb{C}$ as a model for Ω (see Fig. 1), where the embedding $\imath_1$ is given in polar coordinates by $(\varrho, \varphi) \mapsto (\varrho^{2/3}, \frac{2}{3}(\varphi + \pi))$, $\frac{\pi}{2} < \varphi < \frac{3\pi}{2}$ and the other two embeddings are defined similarly. The curves shown on the figure are of the form $t \mapsto \imath_j(s,t)$. This is our analogue of the "pair of pants" surface.

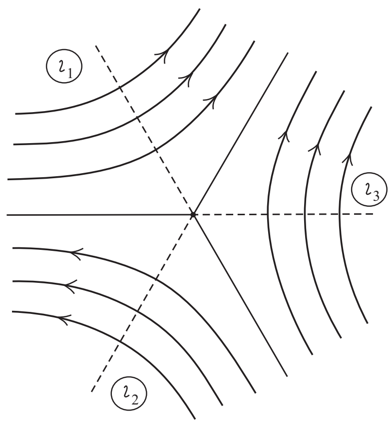

Figure 1: The domain Ω with three long necks.

Remark 2.3. For technical reasons (e.g., to define the Sobolev spaces), it is convenient to endow Ω with a Riemannian metric compatible with the complex structure. Moreover, it is assumed that this metric is the standard flat metric on each of the "long necks" outside a compact subset.

Furthermore, pick any three critical points, say m_1, m_2, m_3 and a pair (γ_1, γ_2) of broken flow lines. More precisely, γ_1 and γ_2 are solutions of the equations

$$\dot\gamma_j + \cos\theta_{j,\nu}(t)\, v_0 + \sin\theta_{j,\nu}(t)\, v_1 = 0,$$
$$\lim_{t\to+\infty} \gamma_j(t) = m_j, \quad \lim_{t\to-\infty} \gamma_j(t) = m_{j+1}.$$

Here $\theta_{j,\nu}(t)$ is a perturbation of the function obtained from $\theta_0(t)$ by putting $(\theta_-, \theta_+) = (\arg(z_j - z_0),\ \arg(z_{j+1} - z_0))$. Then $\mu^2(\gamma_1, \gamma_2)$ should be a formal linear combination of broken flow lines connecting m_1 with m_3. Pick any such flow line, i.e., a solution of the problem[1]

$$\begin{gathered}\dot\gamma_3+\cos\theta_{3,\nu}(t)\,v_0+\sin\theta_{3,\nu}(t)\,v_1=0,\\ \lim_{t\to+\infty}\gamma_3(t)=m_1,\qquad \lim_{t\to-\infty}\gamma_3(t)=m_3,\end{gathered}$$

and also choose $\eta\in\Omega^{0,1}(\Omega)$ such that for $j=1,2,3$ we have $\imath_j^*\eta=\frac12 e^{i\theta_{j,\nu}(t)}d\bar z$ provided $\left|\operatorname{Re} z\right|\ge 1$. Then the multiplicity of γ_3 can conjecturally be defined by counting solutions of the equations

$$\bar\partial u+\eta\otimes v_0(u)=0,\qquad u\colon\Omega\to M, \tag{2.10}$$

$$\lim_{t\to\pm\infty}u\circ\imath_j(s,t)=m_{\sigma_\pm(j)},\qquad \lim_{t\to\pm\infty}\int_0^\infty\left|\partial_s\big(u\circ\imath_j(s,t)\big)\right|ds=0, \tag{2.11}$$

$$\lim_{s\to\infty}u\circ\imath_j(s,t)=\gamma_j(t)\qquad \lim_{s\to\infty}\int_a^b\left|\partial_s\big(u\circ\imath_j(s,t)\big)\right|dt=0. \tag{2.12}$$

Here $\eta\otimes v_0(u)\in\Omega^{0,1}(\Omega;u^*TM)$, $j=1,2,3$, $\sigma_+(1,2,3)=(1,2,1)$, $\sigma_-(1,2,3)=(2,3,3)$. Moreover, in (2.12) "$s\to\infty$" means $s\to-\infty$ for $j=1,2$ and $s\to+\infty$ for $j=3$; The meaning of "∞" in (2.11) is similar.

Notice that over the long necks the above equations and Eqs. (2.7)–(2.9) are of a similar form.

The analogue of holomorphic discs with $d+1$ punctures on the boundary involved in the definition of μ^d are defined in a similar manner.

Remark 2.4. Let $(w_1,\dots,w_{d+1})$ be a tuple of points on the boundary of the unit disc centered at the origin. The indexing is assumed to respect the cyclic ordering, which is obtained by traveling along the circle counterclockwise. Then one can construct a Riemann surface with $(d+1)$ long necks just like Ω (the j's long neck correspond to the sector bounded by l_j and l_{j+1}, where l_j is the ray from the origin containing w_j). Given a $(d+1)$-leafed tree, one can construct another Riemann surface with $d+1$ long necks by gluing the basic ones along long necks. This is similar to [25, 9e]. By stretching interior edges to infinity one obtains the boundary of the moduli space of the Riemann surfaces used in the definition of μ^d.

Let us briefly summarize. We can conjecturally associate with (f,J) a directed A_∞-category $\mathcal A(f,J)$ as follows. The objects of $\mathcal A(f,J)$ are critical

---

[1] Our convention is that for $m_-<m_+$ a broken flow line goes from m_+ to m_- as t varies between $-\infty$ and $+\infty$ and therefore the asymmetry between γ_3 and γ_1,γ_2.

points of f. For any pair (m_-, m_+) of critical points, denote by $CF(m_-, m_+)$ the vector space generated by $\Gamma_\nu(m_-; m_+)$ and put

$$hom_{\mathcal{A}(f,J)}(m_-, m_+) = \begin{cases} CF(m_-, m_+) & m_- < m_+, \\ \mathbb{Z}/2 \cdot id & m_- = m_+, \\ 0 & m_- > m_+. \end{cases}$$

For $\gamma_\pm \in \Gamma_\nu(m_-; m_+)$ denote by $\mathcal{M}^0_\nu(\gamma_-, \gamma_+)$ the zero-dimensional component of the space $\{u \mid u \text{ solves } (2.7)\text{–}(2.9)\}/\mathbb{R}$. Assuming $\#\mathcal{M}^0_\nu(\gamma_-, \gamma_+)$ makes sense, we can define μ^1 by declaring

$$\mu^1(\gamma_-) = \sum_{\gamma_+} \big(\#\mathcal{M}^0_\nu(\gamma_-, \gamma_+) \bmod 2\big)\, \gamma_+.$$

The maps μ^d for $d \geq 2$ are defined in a similar manner and together with μ^1 (conjecturally) combine to an A_∞-structure. Clearly, $\mathcal{A}(f, J)$ depends on the various choices involved in the construction. However, as explained in [24] the derived category $D^b(\mathcal{A}(f, J))$ should not depend on these choices. Moreover, assume (f_τ, J_τ), $\tau \in [0, 1]$ is a continuous family such that f_τ is a J_τ-holomorphic function, whose critical points lie in pairwise different fibres for all τ. Then $D^b(\mathcal{A}(f_0, J_0))$ is equivalent to $D^b(\mathcal{A}(f_1, J_1))$.

Remark 2.5. Our main example is the complex Chern-Simons functional, which takes values in $\mathbb{C}/\mathbb{Z}$ rather than in $\mathbb{C}$. In this case, the construction outlined above does not immediately apply. However, we may proceed as follows. Assume that each line $\ell_r = \{z \mid \operatorname{Re} z = r \mod \mathbb{Z}\}$ contains at most one critical value of f (possibly after a perturbation). Pick r such that the line ℓ_r does not contain any critical value of f and "cut" the cylinder $\mathbb{C}/\mathbb{Z}$ along ℓ_r to obtain a holomorphic function f_r with values in $(0, 1) \times \mathbb{R}$. In other words, consider only those flow lines γ of f for which the image of $f \circ \gamma$ does not intersect the line ℓ_r. Then $D^b(\mathcal{A}(f_r))$ does not depend on r as long as r varies in a connected interval I such that $I \times \mathbb{R}$ does not contain any critical value of f. In this way we obtain a collection of k triangulated categories $\big(D^b(\mathcal{A}(f_{r_j}))\big)_{j=1}^k$, which is well-defined up to a cyclic permutation. Here k is the number of critical values of f.

2.3. A priori C^0-estimates

Since M is not compact, we need to show that solutions of (2.7)–(2.9) do not leave a fixed compact subset of M. This is proved in this subsection under an additional assumption.

The proof of Theorem 2.7, which is the main result of this subsection, crucially depends on the choice of the perturbation θ_ν of the function (2.4). So we take a moment to describe the missing details.

Just like in the beginning of the previous subsection fix a pair of critical points (m_-, m_+) and put $z_\pm = f(m_\pm)$. Up to a translation and a rotation we can assume that

$$z_0 = 0, \quad \theta_\pm \in (0, \pi), \quad \operatorname{Im} z_- = \operatorname{Im} z_+ = \zeta > 0. \tag{2.13}$$

For $\nu \in (0, 1)$ consider a smooth function $\theta_\nu \colon \mathbb{R} \to \mathbb{R}$, which satisfies

$$\theta_\nu(t) = \begin{cases} 0 & |t| \geq \nu^{-1} + 1, \\ \theta_+ & t \in [-\nu^{-1}, -\nu] \\ \theta_- - \pi & t \in [\nu, \nu^{-1}] \end{cases}$$

and is monotone on the intervals $(-\nu^{-1} - 1, -\nu^{-1})$, $(-\nu, \nu)$, and $(\nu^{-1}, \nu^{-1} + 1)$. We also assume that $\theta_\nu(t) \geq 0$ for $t \leq 0$ and that $\theta_\nu(t) \leq 0$ for $t \geq 0$. The graph of θ_ν is shown on Fig. 2.

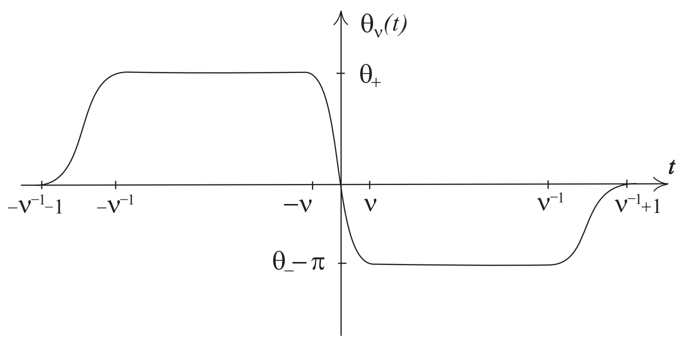

Figure 2: Graph of θ_ν.

Proposition 2.6. *Suppose the closed domain G bounded by the triangle $z_- z_0 z_+$ contains no critical values of f other than $z_\pm$. Let $\nu_j \in (0, 1)$ and $\gamma_j \in \Gamma_{\nu_j}(m_-, m_+)$ be arbitrary sequences such that $\nu_j \to 0$. Then there exists a subsequence $j_k \to \infty$ such that γ_{j_k} converges in $C^0(\mathbb{R}; M)$ and $\gamma_0 = \lim_{k\to\infty} \gamma_{j_k}$ is a solution of (2.5).*

The proof of Proposition 2.6 is given in Appendix B. The assumption that G contains no critical values of f other than $z_\pm$ is essential. If this assumption does not hold, we may have solutions of (2.6), such that the curve $f \circ \gamma_\nu$ is not even homotopic (relative endpoints) to $\ell = \overline{z_+ z_0} \cup \overline{z_0 z_-}$ in $\mathbb{C} \setminus \{z_1, \dots, z_m\}$ (see also Lemma B.1). In particular, the conclusion of Proposition 2.6 is false in this case. To exclude such phenomena, we make the following additional assumption.

(H1) *Convex position of critical values.* The critical values of f are in convex position, i.e., none of the critical values of f is contained in the convex hull of the other critical values. The base point z_0 lies in the interior of the convex hull of the critical values.

Theorem 2.7. *There exists a compact subset $\hat K \subset M \setminus \partial M$ such that the image of any solution u of (2.7) satisfying*

$$\lim_{t \to \pm\infty} u(s,t) = m_\mp, \qquad \lim_{s \to \pm\infty} u(s,t) = \gamma_\mp(t) \tag{2.14}$$

is contained in $\hat K$.

Proof. The proof consists of the following two steps.

Step 1. There exists a constant $\hat R > 0$ such that for any solution $u \in C^2(\mathbb{R}^2)$ of (2.7) and (2.14) we have

$$\bigl| f \circ u(s,t) \bigr| \le \hat R \qquad \textit{for all } (s,t) \in \mathbb{R}^2.$$

Denote $f \circ u = \varphi + i\psi$ and observe that Floer's equation for u implies the equations

$$\partial_s \varphi - \partial_t \psi = \sin\theta_\nu(t)\, \rho \circ u, \quad \partial_s \psi + \partial_t \varphi = -\cos\theta_\nu(t)\, \rho \circ u. \tag{2.15}$$

Denote

$$\Theta_1(t) = \frac{1}{r}\Bigl(\int_0^t \cos\theta_\nu(\tau)\, d\tau - t \Bigr), \qquad \Theta_2(t) = \frac{1}{r} \int_0^t \sin\theta_\nu(\tau)\, d\tau$$

and notice that Θ_1 and Θ_2 are bounded both from above and below (in fact, $\Theta_i(t)$ is locally constant for $|t| \ge \nu^{-1} + 1$). This crucial property is a corollary of our particular choice of θ_ν.

Put $\overline{\Theta}_j = \sup_{\mathbb{R}} \Theta_j(t)$, $\underline{\Theta}_j = \inf_{\mathbb{R}} \Theta_j(t)$, $j = 1, 2$. Furthermore, choose $R > 0$ so large that $f(K) \subset B_R(0)$, where K is the compact subset in (P3). We claim that the following inequality

$$\sup_{\mathbb{R}^2}\big(\varphi(s,t)+\Theta_1(t)\big)\le R+\overline{\Theta}_1 \tag{2.16}$$

holds for all $(s,t)\in\mathbb{R}^2$. We argue by contradiction. Indeed, assume $\varphi(s_0,t_0) + \Theta_1(t_0) = \sup\big(\varphi(s,t)+\Theta_1(t)\big) > R+\overline{\Theta}_1$ for some $(s_0,t_0)\in\mathbb{R}^2$ (the boundary conditions for u imply that the supremum must be attained at some point in $\mathbb{R}^2$). Then $\varphi(s_0,t_0)>R$ so that $(\varphi,\psi)\in B_R^c(0)$ for all (s,t) lying in some small ball B_δ centered at (s_0,t_0). Since $\rho=r^{-1}$ everywhere on $f^{-1}(B_R^c(0))$, from (2.15) we obtain

$$\Delta\varphi = r^{-1}\theta_\nu'(t)\sin\theta_\nu(t),\qquad (s,t)\in B_\delta.$$

Hence, the function $\varphi+\Theta_1$ is harmonic in B_δ and achieves its maximum at an interior point. This contradiction proves (2.16).

Inequality (2.16) implies in turn the estimate

$$\sup_{\mathbb{R}^2}\varphi(s,t)\le R+(\overline{\Theta}_1-\underline{\Theta}_1).$$

Arguing along similar lines one also obtains

$$\inf_{\mathbb{R}^2}\varphi(s,t)\ge -R-(\overline{\Theta}_1-\underline{\Theta}_1),$$

$$\sup_{\mathbb{R}^2}\psi(s,t)\le R+(\overline{\Theta}_2-\underline{\Theta}_2),\qquad \inf_{\mathbb{R}^2}\psi(s,t)\ge -R-(\overline{\Theta}_2-\underline{\Theta}_2).$$

This finishes the proof of Step 1.

Step 2. We prove the theorem.

Let $W\supset\partial M$, $W'\supset\partial M_0\times\mathbb{C}$, and $\psi\colon W'\to W$ be as in the paragraph following (P3). Observe that property (P2) implies that the boundary of M_0 is J_{M_0}-convex, i.e., there exists a function $h\colon M_0\to(-\infty,0]$, which is plurisubharmonic in a neighbourhood of the boundary and $\partial M_0=h^{-1}(0)$. Choose $\varepsilon>0$ so small that h is subharmonic on $h^{-1}(-\varepsilon,0)$ and $U'=h^{-1}(-\varepsilon, 0)\times B_R(0)$ is contained in W'. Denote $U=\psi(U')$.

We claim that for any solution u of (2.7) and (2.14) we have $u(\mathbb{R}^2)\cap U=\varnothing$. Indeed, assuming the converse, there exists $z_0=(s_0,t_0)$ such that $h\circ u(z_0)=\sup\{h\circ u(z)\mid u(z)\in U\}$. Then for sufficiently small $\delta>0$ we can think of u as a map $B_\delta(z_0)\to M_0\times\mathbb{C}$. If π_1 denotes the projection to the first components, the map $\pi_1\circ u$ is pseudoholomorphic. Moreover, $h\circ\pi_1\circ u=h\circ u$ has a local maximum at z_0, which is a contradiction.

Thus the image of u is contained in $\hat{K} = f^{-1}(B_R(0)) \setminus U$. It remains to notice that $\hat{K}$ is compact. □

Remark 2.8. We would like to stress that other results in this paper (except those in Appendix B) depend on hypothesis (H1) only through Theorem 2.7. It is quite possible that an a priori C^0-bound can still be proved for a different choice of the perturbation θ_ν, which does not require convex position of the critical values. However at present it is not quite clear how to obtain such an estimate without (H1).

2.4. The action functional and the energy identity

Denote

$$W^{2,2}_{m_-,m_+} = \big\{\gamma \in W^{2,2}_{loc}(\mathbb{R}; M) \mid \exists T > 0 \text{ and } \xi_\pm \in W^{2,2}\big((T,\infty); T_{m_\pm} M\big) \text{ s.t. } \gamma(\pm t) = \exp_{m_\pm} \xi_\pm(t) \text{ for } t > T\big\}.$$

Then the action functional

$$\mathcal{F}(\gamma) = \int_{\mathbb{R}} \gamma^*\lambda + \int_{\mathbb{R}} \operatorname{Im}\big(e^{-i\theta(t)} f \circ \gamma(t)\big)\, dt \tag{2.17}$$

is well-defined as a map $\mathcal{F}\colon W^{2,2}_{m_-,m_+} \to \mathbb{R}$. Indeed the first integral is convergent, since $\gamma^*\lambda \in W^{1,2}(\mathbb{R}) \hookrightarrow L_1(\mathbb{R})$. As for the second integral, the convergence follows from the fact that f is a quadratic function in an appropriate coordinate chart at $m_\pm$. Observe also, that $\mathcal{F}$ is essentailly the standard symplectic action functional with a Hamiltonian perturbation.

Consider the time-dependent vector field

$$v^t = \operatorname{grad}\operatorname{Re}\left(e^{-i\theta_\nu(t)} f\right) = \cos\theta_\nu(t)\, v_0 + \sin\theta_\nu(t)\, v_1.$$

A standard computation shows that $d\mathcal{F}(\xi) = -\int_{\mathbb{R}} \omega(\xi, \dot\gamma + v^t)\, dt$, where ξ is a vector field along γ. Here we used the fact, that the symplectic gradient of f_0 is $v_1 = \operatorname{grad} f_1$. Therefore with respect to the L^2-metric we have $\operatorname{grad}\mathcal{F} = J(\dot\gamma + v^t)$. Hence, the critical points of the functional $\mathcal{F}$ are broken flow lines of f connecting m_+ and m_-. Similarly, the antigradient flow lines of $\mathcal{F}$ can be interpreted as solutions of Equation (2.7).

Define the energy of a solution u of (2.7) by

$$E(u) = \frac{1}{2}\int_{\mathbb{R}^2} \big(|\partial_s u|^2 + |\partial_t u + v^t|^2\big) ds \wedge dt = \int_{\mathbb{R}^2} |\partial_s u|^2\, ds \wedge dt.$$

Theorem 2.9 (Energy identity). *Let* $u \in C^1(\mathbb{R}^2; M)$ *be a solution of Equations* (2.7)–(2.9). *Then*

$$E(u) = \mathcal{F}(\gamma_+) - \mathcal{F}(\gamma_-).$$

In particular, $E(u) < \infty$.

Proof. It is convenient to denote $\beta_t(s) = u(s,t) = \gamma_s(t)$. Pick arbitrary positive numbers σ and τ. Using Stokes' theorem and the identity

$$\omega(v^t, \partial_s u) = \frac{\partial}{\partial s}\mathrm{Im}\left(e^{-i\theta(t)} f \circ u(s,t)\right)$$

a standard computation yields

$$\begin{aligned}\int_{-\tau}^{\tau}\int_{-\sigma}^{\sigma} |\partial_s u|^2\, ds\wedge dt &= \int_{-\tau}^{\tau}\int_{-\sigma}^{\sigma} \omega\big(\partial_s u, \partial_t u + v^t\big)\, ds\wedge dt \\ &= \int_{-\tau}^{\tau} \gamma_{-\sigma}^*\lambda - \int_{-\tau}^{\tau} \gamma_{\sigma}^*\lambda + \int_{-\sigma}^{\sigma} \beta_{\tau}^*\lambda - \int_{-\sigma}^{\sigma} \beta_{-\tau}^*\lambda \\ &\quad - \int_{-\tau}^{\tau} \mathrm{Im}\, e^{-i\theta(t)} f\circ\gamma_\sigma(t)\, dt + \int_{-\tau}^{\tau} \mathrm{Im}\, e^{-i\theta(t)} f\circ\gamma_{-\sigma}(t)\, dt. \end{aligned} \tag{2.18}$$

With the help of Equation (2.7) we obtain

$$\lambda(\partial_t u) - \lambda(\dot\gamma_\pm) = \lambda(J\partial_s u) + \lambda\big(v^t(\gamma_\pm)\big) - \lambda\big(v^t(u)\big).$$

This in turn implies by (2.9) that

$$\int_{-\tau}^{\tau} \gamma_{\pm\sigma}^*\lambda \longrightarrow \int_{-\tau}^{\tau} \gamma_{\mp}^*\lambda \quad \text{as } \sigma \to +\infty.$$

Similarly, by (2.8) we also have

$$\int_{-\infty}^{+\infty} \beta_{\pm\tau}^*\lambda \longrightarrow 0, \quad \text{as } \tau \to +\infty.$$

Hence, passing in (2.18) first to the limit as $\sigma \to +\infty$ and then to the limit as $\tau \to +\infty$ we obtain the statement of the theorem. □

2.5. A priory C^∞-estimates

It is convenient to introduce the L^p-version of the energy of a map u:

$$E_p(u) = \frac{1}{2}\int_{\mathbb{R}^2} \big(|\partial_s u|^p + |\partial_t u + v^t|^p\big) ds\wedge dt.$$

In particular, $E(u) = E_2(u)$.

Proposition 2.10. *Let u be a solution of* (2.7) *with $E_p(u) < \infty$ for some $p \in [2,\infty)$. Then the following holds:*

$$\lim_{s\to\pm\infty} \partial_s u(s,t) = 0 \qquad \text{and} \qquad \lim_{t\to\pm\infty} \partial_s u(s,t) = 0; \tag{2.19}$$

$$\sup_{\mathbb{R}^2} |\partial_s u| < \infty \qquad \text{and} \qquad \sup_{\mathbb{R}^2} |\partial_t u| < \infty. \tag{2.20}$$

Here both limits in (2.19) *are understood in the $C^0(\mathbb{R})$-topology.*

Proof. The proof is based on the following fact [22, p.12]. There exist constants $h, c > 0$ depending on M, J, ω, and f but not on u such that the following implication holds:

$$\int_{B_r(s,t)} |\partial_s u|^2 < h \quad \Longrightarrow \quad |\partial_s u(s,t)|^2 \le \frac{8}{\pi r^2}\int_{B_r(s,t)} |\partial_s u|^2 + cr^2. \tag{2.21}$$

Pick an arbitrary $\varepsilon \in (0,\varepsilon_0)$, where ε_0 will be chosen below. By the assumption of this proposition there exists $R_\varepsilon > 0$ such that for all $(s,t) \in \mathbb{R}^2$ with $\max\{|s|,|t|\} > R_\varepsilon$ we have $\|\partial_s u\|^2_{L^p(B_1(s,t))} < \varepsilon^{1+\frac{2}{p}}$. Apply the inequality

$$\|\partial_s u\|_{L^2(B_r(s,t))} \le \big(vol(B_r(s,t))\big)^{\frac{1}{2}-\frac{1}{p}} \|\partial_s u\|_{L^p(B_r(s,t))} \tag{2.22}$$

to obtain $\|\partial_s u\|^2_{L^2(B_r(s,t))} \le \pi^{\frac{1}{2}-\frac{1}{p}}\varepsilon^{1+\frac{2}{p}}$ whenever $r \le 1$ and $\max\{|s|,|t|\} > R_\varepsilon$.

Choose $\varepsilon_0 \in (0,1)$ such that $\pi^{\frac{1}{2}-\frac{1}{p}}\varepsilon_0^{1+\frac{2}{p}} < h$. Then (2.21) and (2.22) with $r = \sqrt{\varepsilon}$ yield $|\partial_s u(s,t)|^2 \le (8\pi^{-\frac{2}{p}} + c)\varepsilon$, which proves (2.19).

Furthermore, the first inequality in (2.20) follows immediately from (2.19). The second inequality in (2.20) is obtained from the first one using Equation (2.7) and the fact that $|v^t|^2 = \rho$ is bounded. $\square$

Corollary 2.11. *Let u be a solution of* (2.7) *with $E_p(u) < \infty$ for some $p \in [2,\infty)$. Then the convergence in* (2.14) *in the C^0-topology implies the convergence in the C^1-topology.*

Lemma 2.12. *Let Ω be a bounded domain in $\mathbb{R}^2$. For any integer $k \geq 2$ and any $c_1 > 0$ there exists $c_k = c_k(c_1, \Omega)$ with the following significance. For any solution u of* (2.7) *the following implication holds:*

$$\sup_{\mathbb{R}^2} |\partial_s u| \leq c_1 \qquad \Longrightarrow \qquad \|u\|_{C^k(\Omega)} \leq c_k.$$

The proof of this lemma relies on the local properties of solutions of Floer's equation and can be obtained along the same lines as the proof of Lemma C.3 in [21] (in fact the argument simplifies as we do not need to consider charts with Lagrangian boundary conditions). We omit the details.

Proposition 2.13. *For any integer $k \geq 2$ and any $c_1 > 0$ there exists $c_k = c_k(c_1)$ with the following significance. For any solution u of* (2.7) *the following implication holds:*

$$\sup_{\mathbb{R}^2} |\partial_s u| \leq c_1 \qquad \Longrightarrow \qquad \|u\|_{C^k(\mathbb{R}^2)} \leq c_k.$$

Proof. From Lemma 2.12 we obtain that there exists a constant c_k such that

$$\|u\|_{C^k(\bar{\Omega})} \leq c_k, \tag{2.23}$$

where $\Omega = (0,1) \times (-\nu - 2, \nu + 2)$. This implies that estimate (2.23) is valid for $\Omega = \mathbb{R} \times (-\nu - 2, \nu + 2)$ since Equation (2.7) is invariant with respect to shifts in the s-variable. Applying Lemma 2.12 to $\Omega = (0,1) \times (\nu + 1, \nu + 2)$ and observing that both J and v^t depend neither on s nor on t provided $t \geq \nu + 1$ we obtain that estimate (2.23) also holds for $\Omega = \mathbb{R} \times (\nu + 1, +\infty)$. Similarly, estimate (2.23) is valid for $\Omega = \mathbb{R} \times (-\nu - 1, -\infty)$ as well. This clearly implies the statement of the proposition. □

Theorem 2.14. *Let u be a solution of* (2.7),(2.14). *Assume that for some $p \in [2, \infty)$ there exists a constant $\bar{E}_p$ such that $E_p(u) \leq \bar{E}_p$. Then for any integer $k \geq 0$ there exists a constant $c_k > 0$ such that*

$$\|u\|_{C^k(\mathbb{R}^2)} < c_k.$$

Here constants c_k depend on M, J, f, and $\bar{E}_p$ but not on u.

Proof. It follows from Proposition 2.10 that for any solution u of (2.7) with $E_p(u) < \infty$ we have

$$\|\nabla u\|_{L_\infty} = \max\big\{\sup_{\mathbb{R}^2} |\partial_s u|,\ \sup_{\mathbb{R}^2} |\partial_t u|\big\} < \infty.$$

By the inequality (2.22) we obtain a uniform bound on $\|\partial_s u\|_{L^2(B_1(s,t))}$, where $(s,t) \in \mathbb{R}^2$ is arbitrary. This in turn implies that

$$c_1 = \sup\big\{\|\nabla u\|_{L_\infty} \mid u \text{ solves (2.7),(2.14) and } E_p(u) \le \bar{E}_p\big\} < \infty.$$

Indeed, assuming $c_1 = \infty$ a standard argument [22, p.135] shows that there must be a non-constant holomorphic sphere in M with bounded energy. This is impossible due to the exactness of ω.

The rest follows immediately from Proposition 2.13. □

Recalling Theorem 2.9 we obtain the following result.

Corollary 2.15. *For any integer $k \ge 0$ there exists a constant $c_k > 0$ such that for any solution u of* (2.7)–(2.9) *we have*

$$\|u\|_{C^k(\mathbb{R}^2)} < c_k.$$

2.6. Asymptotic behaviour

Pick any smooth curve $\gamma\colon \mathbb{R} \to M$ such that $\gamma(t) \to m_\pm$ as $t \to \mp\infty$ and denote

$$\sigma(\gamma) = \sigma_\nu(\gamma) = \dot\gamma + v^t(\gamma) \in \Gamma(\gamma^* TM). \tag{2.24}$$

Obviously, $\sigma(\gamma) = 0$ if and only if γ is a broken flow line of f. Consider the linearisation of σ at the point γ:

$$D_\gamma\sigma(\eta) = \nabla_t\,\eta + \nabla_\eta v^t, \qquad \eta \in \Gamma(\gamma^*TM).$$

From now on we assume that all broken flow lines of f are generic. To be more precise, we assume that the following hypothesis holds.

(H2) *Nondegeneracy of broken flow lines.* All solutions of (2.6) are nondegenerate in the following sense: The operator

$$D_\gamma\sigma\colon W^{1,2}(\gamma^*TM) \longrightarrow L^2(\gamma^*TM) \tag{2.25}$$

is an isomorphism.

Remark 2.16. It is proved in Appendix B that hypothesis (H2) holds provided the vanishing cycles corresponding to the segments $\overline{z_0 z_\pm}$ intersect transversely in M_0.

Similarly, pick a smooth map $u\colon \mathbb{R}^2 \to M$ satisfying boundary conditions (2.14) and denote $\Sigma(u) = \partial_s u + J\big(\partial_t u + v^t(u)\big) \in \Gamma(u^*TM)$. Consider the linearisation of Σ at the point u:

$$
\begin{aligned}
D_u\Sigma(\xi) &= \nabla_s \xi + J\big(\nabla_t \xi + \nabla_\xi v^t\big) + \nabla_\xi J(\partial_t u + v^t),\\
&= \nabla_s \xi + J\nabla_t \xi + \cos\theta_\nu \nabla_\xi v_1 - \sin\theta_\nu \nabla_\xi v_0 + \nabla_\xi J(\partial_t u),
\end{aligned}
\tag{2.26}
$$

where $\xi \in \Gamma(\mathbb{R}^2; u^*TM)$.

Remark 2.17. The maps σ and Σ can be viewed as sections of certain Banach bundles (see pp.203 and 186 for details). However this is not needed for the purposes of this subsection.

It is convenient to choose a unitary trivialization Ψ of u^*TM. Recall that for each $(s,t) \in \mathbb{R}^2$ the map $\Psi(s,t)\colon \mathbb{R}^{2n} \to T_{u(s,t)}M$ is a linear isomorphism of complex Hermitian vector spaces, where $\mathbb{R}^{2n}$ is considered to be equipped with the standard complex structure and the standard symplectic form:

$$
J_0 = \begin{pmatrix} 0 & -\mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}, \qquad \omega_0(\xi,\eta) = \xi^t J_0 \eta, \quad \xi,\eta \in \mathbb{R}^{2n}.
$$

Also denote by $\psi_\pm$ the restriction of Ψ to $\gamma_\pm$.

Remark 2.18. One such trivialization Ψ can be constructed as follows. Choose a basis of $T_{m_-}M$ and trivialise γ_-^*TM with the help of the parallel transport along γ_-. Then trivialise u^*TM by doing parallel transport along the curves $\beta_t(\cdot) = u(\cdot,t)$.

With the help of the trivialisations chosen above we can think of the operators $D_{\gamma_\pm}\sigma$ and $D_u\Sigma$ as acting on vector-valued functions. More precisely, there exist matrix-valued functions $S(s,t)$ and $S_\pm(t)$ such that

$$
\begin{aligned}
&\Psi(s,t)\big(\partial_s\xi + J_0\partial_t\xi + S(s,t)\xi\big) = D_u\Sigma\big(\Psi(s,t)\xi\big), &&\forall \xi \in C^\infty(\mathbb{R}^2;\mathbb{R}^{2n});\\
&\psi_\pm(t)\big(\dot\eta - J_0 S_\pm(t)\eta\big) = D_{\gamma_\pm}\sigma\big(\psi_\pm(t)\eta\big), &&\forall \eta \in C^\infty(\mathbb{R};\mathbb{R}^{2n}).
\end{aligned}
$$

Explicitly, matrices S and $S_\pm$ are given by the relations

$$
\Psi(s,t)S(s,t) = \nabla_s\Psi + J\big(\nabla_t\Psi + \nabla_\Psi v^t\big) + \nabla_\Psi J(\partial_t u + v^t), \tag{2.27}
$$

$$
\psi_\pm(t)S_\pm(t) = J\big(\nabla_t\psi_\pm + \nabla_{\psi_\pm} v^t\big). \tag{2.28}
$$

To simplify the notations, denote also by L and $l_\pm$ the operators representing $D_u\Sigma$ and $D_{\gamma_\pm}\sigma$ with respect to the chosen trivialisation:

$$L = \partial_s + J_0\partial_t + S(s,t), \qquad l_\pm = \frac{d}{dt} - J_0 S_\pm(t). \tag{2.29}$$

Lemma 2.19. *Assume the following holds:*

(i) $S\colon \mathbb{R}^2 \to M_{2n}(\mathbb{R})$ *is* C^∞*-bounded*[2];

(ii) $S(s,t)$ *converges to* $S_\pm(t)$ *in the* $C^0(\mathbb{R})$*-topology as* $s \to \mp\infty$;

(iii) *The operators* $l_\pm\colon W^{1,2}(\mathbb{R};\mathbb{R}^{2n}) \to L^2(\mathbb{R};\mathbb{R}^{2n})$ *are invertible;*

(iv) $\lim_{s\to\pm\infty} \sup_t \|\partial_s S(s,t)\| = 0$.

Let ξ *be a solution of the equation* $L\xi = 0$. *If* $\xi \in L^p(\mathbb{R}^2;\mathbb{R}^{2n})$ *for some* $p \in (1,+\infty)$, *then there exist positive constants* C *and* δ *such that*

$$|\xi(s,t)| \le Ce^{-\delta|s|} \qquad \textit{for all } (s,t) \in \mathbb{R}^2.$$

Proof. Let $\xi \in L^p(\mathbb{R}^2;\mathbb{R}^{2n})$ be a solution of the equation $L\xi = 0$. Since L is C^∞-bounded and uniformly elliptic, ξ belongs to $W^{k,\hat p}(\mathbb{R}^2;\mathbb{R}^{2n})$ for all k and $\hat p \in (1,+\infty)$ [26]. In particular, ξ is smooth and for any $s \in \mathbb{R}$ the function $\xi(s,\cdot)$ belongs to $W^{k,2}(\mathbb{R};\mathbb{R}^{2n})$ for all k. The rest of the proof is obtained by applying similar arguments to those used in the proof of Lemma 2.11 in [23]. For the reader's convenience we repeat the main steps here.

Define

$$\chi(s) = \frac{1}{2}\int_{-\infty}^{+\infty} |\xi(s,t)|^2 dt.$$

Then there exists $s_0 \gg 1$ such that for all s with $|s| \ge s_0$ we have

$$\begin{aligned}\chi''(s) &= \int_{-\infty}^{+\infty} \Big(|\partial_s\xi|^2 + \langle \xi, \partial^2_{ss}\xi\rangle\Big)\, dt = 2\int_{-\infty}^{+\infty} |\partial_s\xi|^2 dt + \int_{-\infty}^{+\infty} \langle \xi, (\partial_s S)\xi\rangle\, dt \\ &\ge 2\int_{-\infty}^{+\infty} |J_0\partial_t\xi + S\xi|^2 dt - \varepsilon\int_{-\infty}^{+\infty} |\xi|^2 dt \ge \delta^2 \int_{-\infty}^{+\infty} |\xi|^2 dt = \delta^2\chi(s).\end{aligned}$$

[2]this means that for any integers $\alpha, \beta \ge 0$ there exists $C_{\alpha,\beta} < \infty$ such that $\sup_{\mathbb{R}^2} |\partial_s^\alpha \partial_t^\beta S| \le C_{\alpha,\beta}$

Here we have used the fact that the operator $J_0\partial_t + S(s,t)$ is invertible for $|s| \geq s_0$ and also the following equality:

$$\int_{-\infty}^{+\infty} \langle \xi, \partial_t(J_0\partial_s\xi)\rangle \, dt = \int_{-\infty}^{+\infty} \partial_t\langle \xi, J_0\partial_s\xi\rangle \, dt - \int_{-\infty}^{+\infty} \langle \partial_t\xi, J_0\partial_s\xi\rangle \, dt$$
$$= 0 - \int_{-\infty}^{+\infty} \langle J_0\partial_s\xi + J_0 S\xi, J_0\partial_s\xi\rangle \, dt$$
$$= - \int_{-\infty}^{+\infty} |\partial_s\xi|^2 dt - \int_{-\infty}^{+\infty} \langle S\xi, \partial_s\xi\rangle \, dt.$$

The inequality $\chi''(s) \geq \delta^2\chi(s)$ implies $\chi(s) \leq C_1 e^{-\delta|s|}$. On the other hand, there exists a constant C_2 such that for all solutions of the equation $L\xi = 0$ we have [23] the estimate

$$\Delta|\xi|^2 \geq -C_2|\xi|^2.$$

This implies the mean value inequality

$$|\xi(s,t)|^2 \leq \frac{C_3}{r^2}\int_{B_r(s,t)} |\xi|^2 ds dt.$$

Taking into account the exponential decay of χ we obtain the statement of this lemma from the last inequality. □

Lemma 2.20. *Assume u is a solution of* (2.7),(2.14). *Then there exist positive constants $C_\pm$ such that the estimates*

$$\|S(s,t) - S_\pm(t)\| \leq C_\pm \max\bigl\{|\partial_s u(s,t)|,\, d\bigl(u(s,t), \gamma_\pm(t)\bigr)\bigr\} \tag{2.30}$$

hold for all t and all s such that $\mp s \geq 0$.

Proof. With the help of Equations (2.5) and (2.7) we obtain

$$\begin{aligned}
&\bigl|\Psi(s,t)^{-1}\partial_t u(s,t) - \psi_\pm^{-1}(t)\dot{\gamma}_\pm(t)\bigr| \\
&\leq \bigl|\Psi(s,t)^{-1}(\partial_t u(s,t) + v^t(u(s,t))\bigr| + \bigl|\Psi(s,t)^{-1}v^t(u(s,t)) - \psi_\pm^{-1}(t)v^t(\gamma_\pm(t))\bigr| \\
&\leq \tilde{C}_\pm \max\bigl\{|\partial_s u(s,t)|,\, d\bigl(u(s,t), \gamma_\pm(t)\bigr)\bigr\}
\end{aligned}$$

for some positive constants $\tilde{C}_\pm$ and for all t, s as in the statement of the Lemma. Estimate (2.30) then follows from Formulae (2.27),(2.28), and the above inequality. □

Theorem 2.21 (Exponential decay). *Let u be a solution of* (2.7),(2.14) *with $E_p(u) < \infty$ for some $p \in [2, \infty)$. Then the following holds:*

(i) $\partial_s u \in W^{k,\hat{p}}(\mathbb{R}^2; u^*TM)$ *for all k and all $\hat{p} \in (1, \infty)$. In particular, $E_{\hat{p}}(u) < \infty$ for all $\hat{p} \in (1, \infty)$.*

(ii) *There exist positive constants C, δ such that the inequality*

$$|\partial_s u(s,t)| \le Ce^{-\delta|s|}$$

holds for any $(s,t) \in \mathbb{R}^2$.

Proof. First observe that $\partial_s u$ satisfies $D_u\Sigma(\partial_s u) = 0$ since Equation (2.7) is translation-invariant with respect to the s-variable. Furthermore, we claim that the operator L representing $D_u\Sigma$ in the trivialization Ψ is C^∞-bounded. Indeed, since u is C^∞-bounded, so is $S(s,t)$. Obviuosly, L is also uniformly elliptic and therefore statement *(i)* follows by [26].

To prove *(ii)* it is enough to prove that the matrix-valued function $S(s,t)$ defined by (2.27) satisfies the hypotheses of Lemma 2.19. We have already showed that $S(s,t)$ is C^∞-bounded. From Lemma 2.20 and Proposition 2.10 we obtain that hypothesis *(ii)* of Lemma 2.19 is satisfied. Furthermore, hypothesis *(iii)* is satisfied, since $l_\pm$ represents $D_{\gamma_\pm}\sigma$ in the chosen trivialization. Finally, by *(i)* and the Sobolev embedding theorems any solution u of (2.7) with $E_p(u) < \infty$ satisfies

$$\lim_{s\to\pm\infty} \sup_t \Big(|\nabla_s \, \partial_s u| + |\nabla_t \, \partial_s u| \Big) = 0 \qquad \text{and} \qquad \sup_{\mathbb{R}^2} |\nabla_t \, \partial_t u| < \infty,$$

which implies that hypothesis *(iv)* of Lemma 2.19 is also satisfied. This finishes the proof. □

Corollary 2.22. *Let u be a solution of* (2.7),(2.14) *with $E_p(u) < \infty$ for some $p \in [2, \infty)$. Then*

$$\lim_{t\to\pm\infty} \int_{-\infty}^{+\infty} |\partial_s u(s,t)| \, ds = 0, \qquad \lim_{s\to\pm\infty} \int_a^b |\partial_s u(s,t)| \, dt = 0,$$

i.e., u is a solution of (2.7)–(2.9).

Proof. The statement follows from the Sobolev embedding theorems as explained on p. 159.

□

Corollary 2.23. *For any solution u of* (2.7)–(2.9) $\partial_s u \in W^{k,p}(\mathbb{R}^2; u^*TM)$ *for all k and all $p \in [1, \infty)$. Moreover, for each $k \geq 0$ and $p \geq 1$ there exists a constant $C_{k,p}$ independent of u such that*

$$\|\partial_s u\|_{W^{k,p}} \leq C_{k,p}.$$

Proof. First observe that by Theorem 2.9 $\partial_s u$ belongs to $L^2(\mathbb{R}^2; u^*TM)$ and $\|\partial_s u\|_{L^2}$ is bounded by a constant independent of u.

As already mentioned in the proof of Theorem 2.21 the matrix-valued function $S(s,t)$ is C^∞-bounded. Moreover, it follows from Corollary 2.15 that the corresponding bounds can be chosen to be independent of u. Furthermore, the operator L is uniformly elliptic with the corresponding constant also independent of u. For such an operator of order 1 we have the a priori estimate

$$\|\xi\|_{W^{k,2}} \leq C_k\big(\|L\xi\|_{W^{k-1,2}} + \|\xi\|_{L^2}\big), \qquad k \geq 1, \tag{2.31}$$

where the constant C_k does not depend on u (this is seen by examining explicit formulae for a parametrix of L).

Let ξ represent $\partial_s u$ with respect to the trivialization Ψ. As explained in the proof of Theorem 2.21, $L\xi = 0$. Combining this with (2.31) we obtain the statement of the corollary for all $k \geq 1$ and $p = 2$. This special case implies in turn the statement of the corollary in general by the Sobolev embedding theorems. □

2.7. Compactness

Proposition 2.24. *Let u be a solution of* (2.7) *and* (2.8) *with $E_p(u) < \infty$ for some $p \in [2, \infty)$. Then there exist solutions $\gamma_\pm$ of problem* (2.5) *such that*

$$\lim_{s\to\pm\infty} u(s,t) = \gamma_\mp(t) \qquad \text{and} \qquad \lim_{s\to\pm\infty} \int_a^b |\partial_s u(s,t)|\, dt = 0,$$

where the limits on the left hand side are understood in the $C^0(\mathbb{R})$-topology and $a \leq b$ are arbitrary.

Proof. The proof consists of the following three steps.

Step 1. Let $\beta_n \in C^1(\mathbb{R}; M)$ be an arbitrary sequence of curves such that the following holds:

(i) *There exists a compact subset $\hat{K} \subset M$ containing the images of all curves β_n;*

(ii) $\lim\limits_{n\to\infty} \sup\limits_{t\in\mathbb{R}} \left|\dot\beta_n + v^t\right| = 0$*;*

(iii) $\beta_n(t) \to m_{\mp}$ *as $t \to \pm\infty$ uniformly with respect to n.*

Then there exists a subsequence β_{n_k} converging in $C^0(\mathbb{R}; M)$ to a solution of (2.5)*.*

Recall that the function $\rho = |v_0|^2 = |v^t|^2$ is bounded. Hence, it follows from *(ii)* that the sequence β_n is equicontinuous. By the Arzela-Ascoli theorem there exists a subsequence β_{n_k} convergent on any finite interval to some $\gamma \in C^0(\mathbb{R}; M)$. Then $\gamma \in C^1(\mathbb{R}; M)$ and $\dot\gamma + v^t = 0$.

Furthermore, by *(iii)* for any $\varepsilon > 0$ there exits $T_\varepsilon > 0$ such that for all $t \geq T_\varepsilon$ and all n_k we have $d(\beta_{n_k}(t), m_-) \leq \varepsilon$. Then $d(\gamma(t), m_-) = \lim_{k\to\infty} d(\beta_{n_k}(t), m_-) \leq \varepsilon$ provided $t \geq T_\varepsilon$. Hence $\lim_{t\to+\infty} \gamma(t) = m_-$ and similarly $\lim_{t\to-\infty} \gamma(t) = m_+$, i.e., γ is a solution of (2.5). Then

$$\sup_{t\in\mathbb{R}} d(\beta_{n_k}(t), \gamma(t)) \leq \max\left\{ \sup_{t\in[-T_\varepsilon, T_\varepsilon]} d(\beta_{n_k}(t), \gamma(t)),\ 2\varepsilon \right\} \leq 2\varepsilon$$

provided n_k is large enough. This finishes the proof of Step 1.

Step 2. Let u be a solution of (2.7) *and* (2.8) *with $E_p(u) < \infty$ for some $p \in [2, \infty)$. Then for any $\varepsilon > 0$ there exists $\sigma_\varepsilon > 0$ such that*

$$\inf_{\gamma\in\Gamma(m_-;m_+)} \sup_{t\in\mathbb{R}} d\big(u(s,t), \gamma(t)\big) \leq \varepsilon \qquad \textit{provided } |s| \geq \sigma_\varepsilon.$$

Assume the converse. Then there exists a sequence $s_n \to +\infty$ such that for $\beta_n(t) = u(s_n, t)$ we have

$$\sup_{t\in\mathbb{R}} d\big(\beta_n(t), \gamma(t)\big) \geq \varepsilon_0 \qquad \text{for all } \ \gamma \in \Gamma(m_-; m_+). \tag{2.32}$$

By Theorem 2.7 and Proposition 2.10 the sequence β_n satisfies the hypothesis of Step 1 and hence has a convergent subsequence. But this contradicts inequality (2.32).

Step 3. We prove the proposition.

Let u satisfy the hypotheses of the Proposition. Since $\Gamma(m_-;m_+)$ is discrete, by Step 2 the family $u(s,\cdot)$ converges to some $\gamma_\pm \in \Gamma(m_-;m_+)$ in $C^0(\mathbb{R})$ as $s \to \mp\infty$. The rest follows immediately from Proposition 2.10. □

Proposition 2.25. *For any $\varepsilon > 0$ there exists $T > 0$ such that for all solutions u of* (2.7)–(2.9) *the following holds:*

$$(i) \int\limits_{\mathbb{R}\times[T,+\infty)} |\partial_s u|^2\, dsdt < \varepsilon; \qquad (ii) \int\limits_{\mathbb{R}} |\partial_s u(s,t)|\, ds < \varepsilon \ \text{ for all } t \ge T;$$

$$(iii) \int\limits_{\mathbb{R}\times[-\infty,-T]} |\partial_s u|^2\, dsdt < \varepsilon; \qquad (iv) \int\limits_{\mathbb{R}} |\partial_s u(s,t)|\, ds < \varepsilon \ \text{ for all } t \le -T.$$

Proof. By Corollary 2.23 we have the inequality

$$\int_{\mathbb{R}^2} |\partial_s u(s,t)|\, dsdt < C_{0,1}.$$

This implies that for any $\varepsilon > 0$ and any $T > 0$ there exists $\tau \in [T, T + \varepsilon^{-1}C_{0,1}]$ such that the estimate holds:

$$\int_{\mathbb{R}} |\partial_s u(s,\tau)|\, ds < \varepsilon. \tag{2.33}$$

Arguing like in the proof of the energy identity we obtain the equality

$$\int\limits_{\mathbb{R}\times[\tau,+\infty)} |\partial_s u|^2\, ds\, dt = I(\tau) - \int\limits_{\mathbb{R}} \lambda\big(\partial_s u(s,\tau)\big)\, ds,$$

$$I(\tau) = \int\limits_{\tau}^{+\infty} \Big(\lambda(\dot\gamma_+) - \lambda(\dot\gamma_-) + \operatorname{Im} e^{-i\theta(t)}\big(f\circ\gamma_+ - f\circ\gamma_-\big)\Big)\, dt.$$

Pick any $\varepsilon > 0$ and choose $T_0 > 0$ so large that $|I(\tau)| < \varepsilon$ for all $\tau \ge T_0$. Then, as we have shown above, there exists $\tau \in [T_0, T]$ such that estimate (2.33) holds, where $T = T_0 + \varepsilon^{-1}C_{0,1}$. Hence, we obtain

$$\begin{aligned}\int\limits_{\mathbb{R}\times[T,+\infty)} |\partial_s u|^2\, dsdt &\le \int\limits_{\mathbb{R}\times[\tau,+\infty)} |\partial_s u|^2\, dsdt \\ &\le |I(\tau)| + \Lambda \int_{\mathbb{R}} |\partial_s(s,\tau)|\, ds \le \varepsilon + \Lambda\varepsilon,\end{aligned}$$

where the constant Λ depends on λ only. This proves estimate *(i)*.

Let us prove *(ii)*. We choose $T>0$ so that *(i)* holds. Arguing like in the proof of Corollary 2.23 for $T'>T$ we obtain the inequality

$$\|\partial_s u\|_{W^{k,p}(\mathbb{R}\times[T',+\infty))}\le \tilde C_{k,p}\,\varepsilon,$$

where the constant $\tilde C_{k,p}$ does not depend on u. This in turn implies that there exists a constant $\tilde C$ independent of u such that the inequality $\int_{\mathbb{R}}|\partial_s u(s,t)|\,ds < \tilde C\varepsilon$ holds for all $t\ge T$. This finishes the proof of *(ii)*.

The remaining inequalities are proved in a similar manner. □

Theorem 2.26. *Let $u_k\in\mathcal{M}(\gamma_-;\gamma_+)$ be any sequence. Then there exists a subsequence (still denoted by u_k) and subsequences $s_k^j,\ j=1,\dots,l$, such that $u_k(s+s_k^j,t)$ converges with its derivatives uniformly on compact subsets of $\mathbb{R}^2$ to $u^j\in\mathcal{M}(\gamma^{j-1};\gamma^j)$, where $\gamma^0=\gamma_-,\ \gamma^l=\gamma_+$.*

Proof. We follow the line of argument in [22, p.136]. Denote

$$d_0=\frac13\inf\Big\{d\big(\gamma(0),\delta(0)\big)\mid \gamma,\delta\in\mathcal{M}(m_-;m_+),\ \gamma\ne\delta\Big\}.$$

For an arbitrary sequence $u_k\in\mathcal{M}(\gamma_-;\gamma_+)$ put

$$s_k^1=\sup\{s\in\mathbb{R}\mid d\big(u_k(s,0),\gamma_-(0)\big)>d_0\ \}.$$

Notice that by the definition of s_k^1 we have

$$d\big(u_k(s_k^1,0),\gamma_-(0)\big)=d_0 \qquad\text{and}\qquad d\big(u_k(s+s_k^1,0),\gamma_-(0)\big)\le d_0 \tag{2.34}$$

for all $s\ge 0$.

Since the sequence $\sup_{\mathbb{R}^2}\{|\partial_s u_k|,\ |\partial_t u_k|\}$ is bounded, by [22, Lemma 5.2] we obtain that the sequence $u_k(s+s_k^1,t)$ has a subsequence (still denoted by the same letter) uniformly converging with its derivatives to a map $u^1\colon\mathbb{R}^2\to M$ on compact subsets of $\mathbb{R}^2$. Clearly, u^1 is a solution of (2.7) with $E_2(u^1)\le \mathcal{F}(\gamma_+)-\mathcal{F}(\gamma_-)$. Moreover, by Proposition 2.25 *(ii)* for any $\varepsilon>0$ there exists $T>0$ such that for any $a,b\in\mathbb{R},\ a<b$ we have

$$\int_a^b|\partial_s u^1(s,t)|\,ds=\lim_{k\to\infty}\int_a^b|\partial_s u_k(s+s_k^1,t)|\,ds\le\varepsilon \implies \int_{-\infty}^{+\infty}|\partial_s u^1(s,t)|\,ds\le\varepsilon$$

provided $t > T$. This implies that condition (2.8) holds for $u = u^1$. Then, by Proposition 2.24 we obtain that there exist $\gamma^0, \gamma^1 \in \mathcal{M}(m_-; m_+)$ such that

$$\lim_{s\to+\infty} u^1(s,t) = \gamma^0(t), \qquad \lim_{s\to+\infty} \int_a^b |\partial_s u^1(s,t)|\, dt = 0,$$
$$\lim_{s\to-\infty} u^1(s,t) = \gamma^1(t), \qquad \lim_{s\to-\infty} \int_a^b |\partial_s u^1(s,t)|\, dt = 0.$$

On the other hand, from (2.34) we obtain that $d\big(u^1(s,0), \gamma_-(0)\big) \le d_0$ for all $s \ge 0$. Hence $\gamma^0 = \gamma_-$.

We are done if $\gamma^1 = \gamma_+$. If this is not the case we proceed by induction. Having established the existence of the sequences s_k^j such that $u_k(s + s_k^j, t)$ converges to $u^j \in \mathcal{M}(\gamma^{j-1}, \gamma^j)$ for $j = 1, \dots, q$ we choose $s^* < 0$ such that $d\big(u^q(s^*, 0), \gamma^q(0)\big) < d_0$. For k sufficiently large we then have $d\big(u_k(s_k^q + s^*, 0), \gamma^q(0)\big) < d_0$. Define

$$s_k^{q+1} = \inf\big\{ s \le s_k^q + s^* \mid d\big(u_k(\sigma,0), \gamma^q(0)\big) \le d_0 \quad \text{for } s \le \sigma \le s_k^q + s^* \big\}.$$

Passing to a subsequence if necessary we may assume that $u_k(s + s_k^{q+1}, t)$ converges to $u^{q+1} \in \mathcal{M}(\gamma^q, \gamma^{q+1})$ with $\gamma^{q+1} \ne \gamma^q$. This finishes the induction step. Finally, the process is finite, since for all $q = 1, \dots, l$ we must have $\mathcal{F}(\gamma^{q-1}) < \mathcal{F}(\gamma^q)$. □

Corollary 2.27. *The space*

$$\check{\mathcal{M}}(m_-; m_+) = \bigcup_{\gamma_\pm \in \mathcal{M}(m_-; m_+)} \mathcal{M}(\gamma_-; \gamma_+)$$

is compact.

2.8. Fredholm property

Let

$$A = \sum_{|\alpha| \le m} a_\alpha(x) \frac{\partial}{\partial x^\alpha}, \qquad x \in \mathbb{R}^n$$

be a uniformly elliptic C^∞-bounded differential operator of order m, where a_α takes values in the space of $l \times l$-matrices. Following [19] we say that $A^g = \sum_{|\alpha|\le m} a_\alpha^g(x) \frac{\partial}{\partial x^\alpha}$ is a *limit operator* of A if for some sequence $g \colon \mathbb{N} \to \mathbb{R}^n$ such that $g_j \to \infty$ we have $a_\alpha(x + g_j) \to a_\alpha^g(x)$ uniformly on all compact subsets of $\mathbb{R}^n$.

The following result, which simplifies the arguments used in earlier versions of this paper, has been communicated to the author by V. Rabinovich.

Lemma 2.28. *If A viewed as a map $W^{m,2}(\mathbb{R}^n;\mathbb{R}^l) \to L^2(\mathbb{R}^n;\mathbb{R}^l)$ is Fredholm, then $A\colon W^{k+m,p}(\mathbb{R}^n;\mathbb{R}^l) \to W^{k,p}(\mathbb{R}^n;\mathbb{R}^l)$ is Fredholm for all $k \in \mathbb{R}$, $p > 1$ and its index depends neither on k nor on p.*

Proof. Let A^g be any limit operator of A. Then by Theorem 5.6 of [18] (see also Theorem 2 of [19]) $A^g\colon W^{m,2}(\mathbb{R}^n;\mathbb{R}^l) \to L^2(\mathbb{R}^n;\mathbb{R}^l)$ is invertible. Observe that A^g is a (pseudo)differential operator with the symbol from Hörmander's class $S^m_{1,0}$. By [3, Theorem 3.2] the inverse $(A^g)^{-1}$ is a pseudodifferential operator with the symbol from $S^{-m}_{1,0}$. Hence, $(A^g)^{-1}\colon W^{k-m,2}(\mathbb{R}^n;\mathbb{R}^l) \to W^{k,2}(\mathbb{R}^n;\mathbb{R}^l)$ is bounded for any $k \in \mathbb{R}$. Applying [18, Theorem 5.6] again we obtain that $A\colon W^{k+m,2}(\mathbb{R}^n;\mathbb{R}^l) \to W^{k,2}(\mathbb{R}^n;\mathbb{R}^l)$ is Fredholm for any $k \in \mathbb{R}$.

For arbitrary k and $p > 1$ put $k' = \min\{k-1, k - \frac{n}{p} + \frac{n}{2}\}$ to obtain the embeddings $W^{k,p}(\mathbb{R}^n;\mathbb{R}^l) \hookrightarrow W^{k',2}(\mathbb{R}^n;\mathbb{R}^l)$, $W^{k+m,p}(\mathbb{R}^n;\mathbb{R}^l) \hookrightarrow W^{k'+m,2}(\mathbb{R}^n;\mathbb{R}^l)$. We claim that $A\colon W^{k+m,p}(\mathbb{R}^n;\mathbb{R}^l) \to W^{k,p}(\mathbb{R}^n;\mathbb{R}^l)$ has a closed range. Indeed, let ζ_n be any sequence from $A\big(W^{k+m,p}(\mathbb{R}^n;\mathbb{R}^l)\big)$ converging to ζ_0 in $W^{k,p}(\mathbb{R}^n;\mathbb{R}^l)$. Then $\zeta_n \in A\big(W^{k'+m,2}(\mathbb{R}^n;\mathbb{R}^l)\big)$ converges to ζ_0 in $W^{k',2}(\mathbb{R}^n;\mathbb{R}^l)$. Hence $\zeta_0 = A\xi_0$ for some $\xi_0 \in W^{k'+m,2}(\mathbb{R}^n;\mathbb{R}^l)$. Since A is C^∞-bounded uniformly elliptic operator, $\zeta_0 \in W^{k,p}(\mathbb{R}^n;\mathbb{R}^l)$ implies that $\xi_0 \in W^{k+m,p}(\mathbb{R}^n;\mathbb{R}^l)$. This proves that $A\big(W^{k+m,p}(\mathbb{R}^n;\mathbb{R}^l)\big)$ is closed in $W^{k,p}(\mathbb{R}^n;\mathbb{R}^l)$.

Furthermore, C^∞-boundedness and uniform ellipticity imply that if $\xi \in W^{k,p}(\mathbb{R}^n;\mathbb{R}^l)$ is in the kernel of A for some k and p, then $\xi \in W^{k,p}(\mathbb{R}^n;\mathbb{R}^l)$ for all k and p. In particular, for any k and p the dimension of $\ker\big(A\colon W^{k+m,p}(\mathbb{R}^n;\mathbb{R}^l) \to W^{k,p}(\mathbb{R}^n;\mathbb{R}^l)\big)$ is finite and depends neither on k nor on p. Moreover, applying similar arguments to the formal adjoint operator of A we obtain that the dimension of $\operatorname{coker}\big(A\colon W^{k+m,p}(\mathbb{R}^n;\mathbb{R}^l) \to W^{k,p}(\mathbb{R}^n;\mathbb{R}^l)\big)$ is also finite and depends neither on k nor on p. □

In the lemma below we use the same notations as in Lemma 2.19.

Lemma 2.29. *Assume that hypotheses (i)–(iii) of Lemma 2.19 as well as the following holds:*

(a) *For each $t \in \mathbb{R}$ the matrix $J_0 S_\pm(t)$ is symmetric;*

(b) *$S(s,t)$ converges to constant matrices $H_\pm$ in the C^0-topology as $t \to \mp\infty$. Moreover,*

$$H_+ = \lim_{t\to-\infty} S_+(t) = \lim_{t\to-\infty} S_-(t), \qquad H_- = \lim_{t\to+\infty} S_+(t) = \lim_{t\to+\infty} S_-(t)$$

are symmetric matrices.

Then $L\colon W^{k+1,p}(\mathbb{R}^2;\mathbb{R}^{2n}) \to W^{k,p}(\mathbb{R}^2;\mathbb{R}^{2n})$ *is Fredholm for any* $k \in \mathbb{R}$, $p > 1$ *and its index depends neither on* k *nor on* p.

Proof. The proof consists of the following three steps.

Step 1. Consider the s*-independent operators*

$$L_\pm = \partial_s + J_0\partial_t + S_\pm(t).$$

Then $L_\pm\colon W^{1,2}(\mathbb{R}^2;\mathbb{R}^{2n}) \to L^2(\mathbb{R}^2;\mathbb{R}^{2n})$ *are invertible.*

First observe that since $l_\pm\colon W^{1,2}(\mathbb{R};\mathbb{R}^{2n}) \to L^2(\mathbb{R};\mathbb{R}^{2n})$ are isomorphisms we have the estimates

$$\|\eta\|_{W^{1,2}} \le C_\pm \|l_\pm \eta\|_{L^2}, \qquad \text{for all } \eta \in W^{1,2}(\mathbb{R};\mathbb{R}^{2n}). \tag{2.35}$$

Let $\hat\chi(\sigma)$ denote the Fourier transform of a function $\chi(s)$. Pick any $\zeta \in C^\infty(\mathbb{R}^2;\mathbb{R}^{2n})$ with compact support and for any fixed t apply the Fourier transform in the variable s to the equation $L_\pm(\eta) = \zeta$ to obtain

$$i\sigma\,\hat\eta(\sigma,t) + J_0 l_\pm \hat\eta(\sigma,t) = \hat\zeta(\sigma,t).$$

Observe that $J_0 l_\pm$ is a symmetric operator such that 0 does not belong to the spectrum of $J_0 l_\pm$. Hence, the above equation is solvable for any real σ. Applying the inverse Fourier transform we obtain a solution η of the initial equation $L_\pm(\eta) = \zeta$. Moreover, with the help of (2.35) an easy computation yields the estimate $\|\eta\|_{W^{1,2}} \le \tilde C_\pm \|\zeta\|_{L_2}$. This implies that $L_\pm$ are isomorphisms.

Step 2. Consider the operators

$$K_\pm = \partial_s + J_0\partial_t + H_\pm$$

with constant coefficients. Then $K_\pm\colon W^{1,2}(\mathbb{R}^2;\mathbb{R}^{2n}) \to L^2(\mathbb{R}^2;\mathbb{R}^{2n})$ *are invertible.*

Write $J_0K_\pm = -\partial_t + J_0\partial_s + J_0H_\pm$ and observe that the operators

$$k_\pm = \tfrac{d}{ds} + H_\pm\colon W^{1,2}(\mathbb{R};\mathbb{R}^{2n}) \to L^2(\mathbb{R};\mathbb{R}^{2n})$$

are isomorphisms. Indeed, any function satisfying $k_\pm\eta = 0$ can be expressed through exponential functions and therefore does not belong to $W^{1,2}(\mathbb{R};\mathbb{R}^{2n})$.

Similarly, the cokernel of $k_\pm$ is also trivial. The rest of the proof of this step is analogous to the proof of Step 1.

Step 3. We prove the lemma.

Clearly, any limit operator[3] $L_0 = \partial_s + J_0\partial_t + S_0(s,t)$ of L must be $K_\pm$ or

$$L^\tau_\pm = \partial_s + J_0\partial_t + S_\pm(t+\tau) = V_\tau L_\pm V_{-\tau},$$

where V_τ denotes the shift operator $\xi(t) \mapsto \xi(t+\tau)$. Since V_τ acts as an isomorphism on $W^{k,2}(\mathbb{R}^2,\mathbb{R}^{2n})$ for any k, the operator $L^\tau_\pm\colon W^{1,2}(\mathbb{R}^2,\mathbb{R}^{2n}) \to L^2(\mathbb{R}^2,\mathbb{R}^{2n})$ is also an isomorphism. Hence, by [18, Theorem 5.6] we obtain that $L\colon W^{1,2}(\mathbb{R}^2,\mathbb{R}^{2n}) \to L^2(\mathbb{R}^2,\mathbb{R}^{2n})$ is Fredholm. Then the statement of the lemma follows from Lemma 2.28. □

Theorem 2.30. *For each solution u of (2.7)–(2.9) the map*

$$D_u\Sigma\colon W^{k+1,p}(\mathbb{R}^2; u^*TM) \to W^{k,p}(\mathbb{R}^2; u^*TM)$$

is Fredholm for any $k \in \mathbb{R}$, $p > 1$ and its index depends neither on k nor on p.

Proof. Clearly, it is enough to check that the matrix-valued function $S(s,t)$ given by (2.27) satisfies the hypotheses of Lemma 2.29. The fact that *(a)* holds can be checked by direct computation using (2.28) and is well known [23]. To see that *(b)* holds, observe that $v^t = v_0$ for $|t|$ large enough. It follows that $H_\pm$ represents $J\nabla v_0 = \nabla v_1$ at $m_\pm$. Here we used the fact, that J is integrable in a neighbourhood of $m_\pm$. It remains to notice that ∇v_1 is the Hessian of $f_1 = \operatorname{Im} f$ at $m_\pm$ and therefore is symmetric. □

To compute the index of $D_u\Sigma$ we need some preparation. Since $\operatorname{Ind}(D_u\Sigma)$ depends neither on k nor on p, we can put $k = 1,\ p = 2$. With an arbitrary C^1-curve $\gamma\colon \mathbb{R} \to M$ satisfying

$$\lim_{t\to\pm\infty} \gamma(t) = m_\mp \quad \text{and} \quad \lim_{t\to\pm\infty} \dot\gamma(t) = 0 \tag{2.36}$$

we associate a pair of Lagrangian subspaces in $T_{\gamma(0)}M$ as follows. Consider the operator

$$A\colon C^\infty(\gamma^*TM) \to C^\infty(\gamma^*TM), \qquad A\xi = J\nabla_t\,\xi + \tilde S\xi,$$

[3] in the sense explained in the beginning of this subsection

where $\tilde{S}$ is a zero-order operator, namely $\tilde{S}\xi = \nabla_\xi \tilde{v}^t + (\nabla_\xi J)\dot{\gamma}$, $\tilde{v}^t = Jv^t = \cos\theta_\nu(t)v_1 - \sin\theta_\nu(t)v_0$ (compare with (2.26)). Notice that $\lim_{t\to\mp\infty} \tilde{S} = \tilde{S}^\pm \in End(T_{m_\pm}M)$ is the Hessian of $\pm\mathrm{Im}\,(e^{-i\theta_\pm}f)$ at $m_\pm$. As we already observed in the proof of Theorem 2.30 $J\tilde{S}^\pm$ is then the Hessian of $\pm\mathrm{Re}\,(e^{-i\theta_\pm}f)$ and therefore is a non-degenerate self-adjoint endomorphism with vanishing signature (i.e., $J\tilde{S}^\pm$ has n positive and n negative eigenvalues).

Denote by ξ_v, $\mathrm{v}\in \mathrm{T}_{\gamma(0)}\mathrm{M}$, a solution of the Cauchy problem $A\xi_\mathrm{v} = 0$, $\xi_\mathrm{v}(0) = \mathrm{v}$ and put

$$\Lambda^\pm = \{\, \mathrm{v}\in \mathrm{T}_{\gamma(0)}\mathrm{M} \;\mid\; \lim_{\mathrm{t}\to\mp\infty} \xi_\mathrm{v}(\mathrm{t}) = 0\}. \tag{2.37}$$

Then $\Lambda^\pm$ are Lagrangian subspaces. Indeed, a straightforward computation shows that $\omega(\xi_\mathrm{v}(t), \xi_\mathrm{w}(t))$ does not depend on t for any $\mathrm{v}, \mathrm{w}\in \mathrm{T}_{\gamma(0)}\mathrm{M}$. Therefore, if $\mathrm{v}, \mathrm{w}\in\Lambda^+$, then $\omega(\mathrm{v},\mathrm{w}) = 0$ since $\omega(\xi_\mathrm{v}(t), \xi_\mathrm{w}(t))$ vanishes at $-\infty$. Besides, $\dim\Lambda^+ = n$ since the signature of $J\tilde{S}^+$ vanishes.

Remark 2.31. If $\mathrm{v}\in\Lambda^\pm$, then ξ_v decays exponentially fast at $\mp\infty$ since $J\hat{S}^\pm$ is nondegenerate and self-adjoint. Hence, the kernel of the operator $A\colon W^{1,2}(\gamma^*TM)\to L^2(\gamma^*TM)$ can be identified with $\Lambda^+\cap\Lambda^-$. In particular, $\ker A$ is nontrivial if and only if $\Lambda^+\cap\Lambda^- \neq \{0\}$.

Furthermore, pick any two curves $\gamma_\pm$ satisfying (2.36) such that the associated pairs of Lagrangian subspaces are transverse. Let $u\colon\mathbb{R}^2\to M$ be any C^1-map such that each curve $\gamma_s(t) = u(s,t)$ also satisfies (2.36) and $\gamma_s\to\gamma_\pm$ as $s\to\mp\infty$ in the C^1-topology. With the help of the relative Maslov index for Lagrangian pairs [20] we associate with the triple $(\gamma^+,\gamma^-;u)$ an integer $\mu(\gamma^+,\gamma^-;u)$, which is referred to as the *relative Maslov index*. To define $\mu(\gamma^+,\gamma^-;u)$ for any fixed s denote by $(\Lambda^+(s),\Lambda^-(s))$ the pair of Lagrangian subspaces of $T_{\gamma_s(0)}M$ as in (2.37) with $\gamma(t) = \gamma_s(t)$. Furthermore, denote by $\mathcal{L}(TM)$ the Lagrangian Grassmannian bundle and put $\beta_0(s) = \gamma_s(0) = u(s,0)$. Then (Λ^+,Λ^-) can be viewed as a pair of sections of the bundle $\beta_0^*\mathcal{L}(TM)$ such that the subspaces $\Lambda^+(s)$ and $\Lambda^-(s)$ are transverse for $s=\pm\infty$. Choose a unitary trivialization of β_0^*TM and represent $\Lambda^\pm$ by a pair of curves $\Lambda_0^\pm\colon\mathbb{R}\to\mathcal{L}(\mathbb{R}^{2n})$. It is said that a crossing, i.e., a point s_0 such that $\Lambda^+(s_0)\cap\Lambda^-(s_0)\neq 0$, is *regular* if the associated (relative) crossing form [20, p.834] $\Gamma(\Lambda^+,\Lambda^-,s_0)\colon\Lambda^+(s_0)\cap\Lambda^-(s_0)\to\mathbb{R}$ is nondegenerate. If all crossings are regular, then the number

$$\mu(\gamma^+,\gamma^-;u) = \mu(\Lambda_0^+,\Lambda_0^-) = \sum_{s_0 \text{ is crossing}} \operatorname{sign}\Gamma(\Lambda^+,\Lambda^-,s_0) \;\in\mathbb{Z}$$

does not depend on the choice of the unitary trivialization, i.e. the relative Maslov index is well-defined.

Proposition 2.32. *With the same notations as in Theorem 2.30, the index of $D_u\Sigma$ is given by*

$$\operatorname{Ind}(D_u\Sigma) = \mu(\gamma^+, \gamma^-; u).$$

Proof. We follow the line of argument in [23].

Choose a C^1-small perturbation $\hat{u}$ of the map u with the following properties: $D_{\hat{u}}\Sigma$ is Fredholm, $\operatorname{Ind}(D_{\hat{u}}\Sigma) = \operatorname{Ind}(D_u\Sigma)$, the Lagrangian pairs associated with the curves $\hat{u}(\pm\infty, t)$ are transverse, and there exists $T > 0$ such that $\hat{u}(s, \pm t) = m_\mp$ for all $t \geq T$. Construct also a unitary trivialization of $\hat{u}^*TM$ as described in Remark 2.18. Write $D_{\hat{u}}\Sigma$ in the form

$$L(\xi) = \partial_s\xi + A(s)\xi, \qquad \xi\colon \mathbb{R}^2 \to \mathbb{R}^{2n},$$

where $A(s)\xi = J_0\partial_t\xi + \hat{S}(s,t)\xi$. Since the limits of the matrix $S(s,t)$ associated with u are symmetric, up to a compact perturbation we can also assume that matrix $\hat{S}(s,t)$ is symmetric for all (s,t). By the choice of $\hat{u}$ we also have $\hat{S}(s, \pm t) = H_\mp$ for $t \geq T$, where $H_\mp$ represents the Hessian of $\operatorname{Im} f$ at $m_\mp$.

Furthermore, denote $\mu_0 = \min\{|\mu| : \ker(J_0H_\pm - \mu) \neq 0\} > 0$ and consider $A(s)$ as an unbounded operator in $L^2(\mathbb{R}; \mathbb{R}^{2n})$ with the domain $W^{1,2}(\mathbb{R}; \mathbb{R}^{2n})$. Then for all $s \in \mathbb{R}$ any point of the spectrum of $A(s)$ from the interval $(-\mu_0, \mu_0)$ is an eigenvalue. Indeed, for any $(s, \mu) \in \mathbb{R} \times (-\mu_0, \mu_0)$ the operator $A(s) - \mu\colon W^{1,2}(\mathbb{R}; \mathbb{R}^{2n}) \to L^2(\mathbb{R}; \mathbb{R}^{2n})$ is Fredholm, since $J_0H_\pm - \mu$ is nondegenerate. Moreover, $\operatorname{Ind}(A(s) - \mu) = \operatorname{Ind} A(s) = \operatorname{Ind}(-J_0A(s)) = 0$ since the indices of $-J_0A(+\infty)$ and $D_{\gamma_-}\sigma$ coincide. Hence, if μ is not an eigenvalue of $A(s)$, then $A(s) - \mu\colon W^{1,2}(\mathbb{R}; \mathbb{R}^{2n}) \to L^2(\mathbb{R}; \mathbb{R}^{2n})$ is bijective and therefore μ belongs to the resolvent set of $A(s)$.

From the above observation follows [1] that the index of L can be computed with the help of the spectral flow of $A(s)$. Namely, a point s_0 is said to be a regular crossing of the family $A(s)$ if $\ker A(s_0) \neq 0$ and the crossing form

$$\Gamma(A, s_0)\xi = \langle \xi, (\partial_s A)\xi \rangle_{L^2} = \langle \xi, \partial_s S(s_0, \cdot)\xi \rangle_{L^2}, \qquad \xi \in \ker A(s_0)$$

is nondegenerate. Then, if $A(s)$ has only regular intersection points, we have: $\operatorname{Ind} L = \sum_{s_0} \operatorname{sign} \Gamma(A, s_0)$, where the summation runs over all crossings s_0.

It follows from Remark 2.31 that crossings of $(\Lambda_0^+, \Lambda_0^-)$ and $A(\cdot)$ coincide. Therefore to complete the proof it suffices to show that under the natural identification $\Lambda_0^+(s_0) \cap \Lambda_0^-(s_0) \cong \ker A(s_0)$ the associated crossing forms coincide at each crossing s_0 (we can assume that only regular crossings occur).

Let $\Xi(s,t)$ be the solution operator of $A(s)$, i.e., $\Xi(s,t)$ is a square matrix of dimension $2n$ satisfying

$$J_0 \partial_t \Xi + S(s,t)\Xi = 0, \quad \Xi(s,0) = \mathbb{1}.$$

Since $S(s,t)$ is symmetric, $\Xi(s,t) \in Sp(2n;\mathbb{R})$ for all (s,t). From the equality

$$\partial_t(\Xi^T J_0 \partial_s \Xi) = -(\Xi^T S J_0) J_0 \partial_s \Xi + \Xi^T J_0 \partial_s (J_0 S \Xi) = -\Xi^T \partial_s S\, \Xi$$

we obtain $\langle \Xi\xi_0,\, \partial_s S \Xi \xi_0 \rangle = -\partial_t \, \langle \Xi\xi_0,\, J_0 \partial_s \Xi \xi_0 \rangle = \partial_t \, \omega_0\big(\Xi\xi_0,\, \partial_s \Xi \xi_0\big)$, where $\xi_0 \in \mathbb{R}^{2n}$. Hence, for any crossing s_0 and any $\xi_0 \in \Lambda_0^+(s_0) \cap \Lambda_0^-(s_0)$ we have:

$$\begin{aligned} \Gamma(A,s_0)\xi_0 &= \int_{-\infty}^{+\infty} \big\langle \Xi(s_0,t)\xi_0,\, \partial_s S(s_0,t)\Xi(s_0,t)\xi_0 \big\rangle \, dt \\ &= \lim_{t\to+\infty} \omega_0\big(\Xi(s_0,t)\xi_0,\, \partial_s \Xi(s_0,t)\xi_0\big) \\ &\quad - \lim_{t\to-\infty} \omega_0\big(\Xi(s_0,t)\xi_0,\, \partial_s \Xi(s_0,t)\xi_0\big). \end{aligned} \tag{2.38}$$

On the other hand, for ξ_0 as above and for all s from a sufficiently small neighbourhood of s_0 there exists $\xi^-(s) \in \Lambda^-(s_0)$ such that $\xi_0 + \xi^-(s) \in \Lambda^+(s)$, i.e.,

$$\lim_{t\to+\infty} \Xi(s_0,t)\xi^-(s) = 0 \quad \text{and} \quad \lim_{t\to-\infty} \Xi(s,t)(\xi_0 + \xi^-(s)) = 0.$$

For $t \le -T$ we must have $\Xi(s,t)(\xi_0 + \xi^-(s)) = \sum_{j=1}^n c_j(s) e^{\lambda_j t}$, where $\lambda_1, \ldots, \lambda_n$ are positive eigenvalues of the matrix $J_0 H_+$. Hence,

$$\partial_s \Xi(s_0,t)\xi_0 + \Xi(s_0,t)\partial_s \xi^-(s_0) \to 0$$

as $t \to -\infty$ and this in turn implies

$$\begin{aligned} \omega_0\big(\xi_0, \partial_s \xi^-(s_0)\big) &= \omega_0\big(\Xi(s_0,t)\xi_0, \Xi(s_0,t)\partial_s \xi^-(s_0)\big) \\ &= -\lim_{t\to-\infty} \omega_0\big(\Xi(s_0,t)\xi_0, \partial_s \Xi(s_0,t)\xi_0\big). \end{aligned} \tag{2.39}$$

Similarly, there also exists $\xi^+(s) \in \Lambda^+(s_0)$ for all s sufficiently close to s_0 such that $\xi_0 + \xi^+(s) \in \Lambda^-(s)$. Arguing as above, wee see that

$$\omega_0\big(\xi_0, \partial_s \xi^+(s_0)\big) = -\lim_{t\to+\infty} \omega_0\big(\Xi(s_0,t)\xi_0, \partial_s \Xi(s_0,t)\xi_0\big). \tag{2.40}$$

Recalling the definition of the crossing form and Theorem 1.1 of [20], we obtain $\Gamma(\Lambda^+,\Lambda^-,s_0)\xi_0 = \omega_0\big(\xi_0,\partial_s\xi^-(s_0)\big) - \omega_0\big(\xi_0,\partial_s\xi^+(s_0)\big)$. Combining this with (2.38)-(2.40) we finally obtain $\Gamma(\Lambda^+,\Lambda^-,s_0)\xi_0 = \Gamma(A,s_0)\xi_0$. This finishes the proof. $\square$

For any $p > 2$ consider the space

$$\begin{aligned}\mathcal{B} = \big\{u \in W^{1,p}_{loc}(\mathbb{R}^2;M) \mid \exists\, R>0, \xi_\pm \in W^{1,p}(\gamma_\pm^* TM) \text{ and}\\ \eta_\pm \in W^{1,p}(\mathbb{R}^2; T_{m_\pm}M) \text{ s.t. } u = \exp_{\gamma_\pm}\xi_\pm\\ \text{for } \mp s > R \text{ and } u = \exp_{m_\pm}\eta_\pm \text{ for } \mp t > R\big\}.\end{aligned}$$

One can construct an atlas on $\mathcal{B}$ similarly to [11, Theorem 3]. Thus $\mathcal{B}$ is a Banach manifold. Observe that for $u\in\mathcal{B}$ we have $T_u\mathcal{B} = W^{1,p}(\mathbb{R}^2;u^*TM)$.

Let $\mathcal{F}\to\mathcal{B}$ be the vector bundle with the fiber $\mathcal{F}_u = L^p(\mathbb{R}^2;u^*TM)$. Then the map Σ can be interpreted as a section of $\mathcal{F}$. Clearly, any solution of $\Sigma(u)=0$ is a smooth map. By Corollary 2.22 the zero locus of Σ coincides with $\mathcal{M}(\gamma_-;\gamma_+)$. Notice also that the covariant derivative of Σ at the point u can be identified with the map

$$D_u\Sigma\colon W^{1,p}(\mathbb{R}^2;u^*TM)\to L^p(\mathbb{R}^2;u^*TM),$$

which is Fredholm. We summarise these observations in the following proposition.

Proposition 2.33. *The zero locus of* $\Sigma\in\Gamma(\mathcal{B};\mathcal{F})$ *is the space of solutions of* (2.7)–(2.9). *Moreover, for each zero* u *the covariant derivative* $D_u\Sigma$ *is Fredholm.*

3. A gauge theory on 5-manifolds

Let E be a five-dimensional oriented Euclidean vector space with a preferred vector $\mathrm{v}\in E$ of unit norm. Let $\eta(\cdot) = \langle \mathrm{v},\cdot\rangle$ denote the corresponding 1-form and $*$ the Hodge operator. Then the linear map

$$T_\eta\colon \Lambda^2E^*\longrightarrow\Lambda^2E^*,\quad \omega\mapsto *(\omega\wedge\eta)$$

has three eigenvalues $\{-1,0,+1\}$ and the space Λ^2E^* decomposes as the direct sum of the corresponding eigenspaces:

$$\Lambda^2E^*\cong\Lambda^2_-E^*\oplus\Lambda^2_0E^*\oplus\Lambda^2_+E^*.$$

Indeed, denote by H the orthogonal complement of v. Then $\Lambda^2 E^* \cong \Lambda^2 H^* \oplus H^*$ and one easily checks that the following subspaces $\Lambda^2_\pm H^*$ and H^* are eigenspaces of T_η, where $\Lambda^2_\pm H^*$ denote the eigenspaces of the 4-dimensional Hodge star operator. In other words, $\Lambda^2_\pm E^* \cong \Lambda^2_\pm H^*$ and $\Lambda^2_0 E^* \cong H^*$.

Identify the Clifford algebra of E with ΛE and recall the following description of the Clifford multiplication

$$
\begin{aligned}
&Cl\colon E^* \otimes \Lambda E^* \longrightarrow \Lambda E^*, && Cl = Cl' + Cl'',\\
&Cl'\colon E^* \otimes \Lambda^p E^* \cong E \otimes \Lambda^p E^* \xrightarrow{\ c\ } \Lambda^{p-1} E^*, && c(e \otimes \omega) = -\imath_e \omega,\\
&Cl''\colon E^* \otimes \Lambda^p E^* \xrightarrow{\ \cdot\wedge\cdot\ } \Lambda^{p+1} E^*.
\end{aligned}
$$

In particular, by restriction we get a map $Cl'\colon E^* \otimes \Lambda^2_+ E^* \longrightarrow E^*$, which is essentially the four-dimensional homomorphism $H^* \otimes \Lambda^2_+ H^* \longrightarrow H^*$.

Observe that $\Lambda^2_+ H^*$ has a natural structure of a Lie algebra as a three-dimensional oriented Euclidean vector space. For an arbitrary Lie algebra $\mathfrak{g}$ denote $V = \Lambda^2_+ H^* \otimes \mathfrak{g}$ and consider the map $\sigma\colon V \otimes V \to V$, $\sigma = \frac{1}{2}[\cdot\,,\cdot]_{\Lambda^2_+ H^*} \otimes [\cdot\,,\cdot]_{\mathfrak{g}}$. Choosing a Lie algebra isomorphism $\Lambda^2_+ H^* \cong \mathbb{R}^3$, for $\xi = e_1 \otimes \xi_1 + e_2 \otimes \xi_2 + e_3 \otimes \xi_3 \in V \cong \mathbb{R}^3 \otimes \mathfrak{g}$ we obtain

$$\sigma(\xi, \xi) = e_1 \otimes [\xi_2, \xi_3] + e_2 \otimes [\xi_3, \xi_1] + e_3 \otimes [\xi_1, \xi_2].$$

Let (W^5, g) be an arbitrary oriented Riemannian five-manifold with a preferred vector field v of pointwise unit norm. Denote $\eta(\cdot) = g(v, \cdot) \in \Omega^1(W)$ and $\mathcal{H} = \ker \eta \subset TW$. As described above, we have the following splittings:

$$
\begin{aligned}
&\Omega^1(W) = \Omega^1_h(W) \oplus \Omega^0(W)\eta, \qquad \Omega^1_h(W) = \Gamma(\mathcal{H}^*),\\
&\Omega^2(W) = \Omega^2_-(W) \oplus \Omega^2_0(W) \oplus \Omega^2_+(W),
\end{aligned}
$$

where $\Gamma(\mathcal{H}^*)$ is the space of sections of $\mathcal{H}^*$. Let $P \to W$ be a principal G-bundle, where G is a compact Lie group. Denote by $\mathcal{A}(P)$ the space of connections on P and by $ad\,P$ the adjoint bundle of Lie algebras. Consider the following equations for a pair $(A, B) \in \mathcal{A}(P) \times \Omega^2_+(ad\,P) = \mathcal{B}$:

$$\imath_v F_A - \delta^+_A B = 0, \qquad F^+_A - \nabla^A_v B - \sigma(B, B) = 0, \tag{3.1}$$

where the operator $\delta^+_A\colon \Omega^2_+(ad\,P) \to \Omega^1_h(ad\,P)$ is defined by the composition

$$
\begin{aligned}
\delta^+_A\colon \Gamma(\Lambda^2_+ \mathcal{H}^* \otimes ad\,P) &\xrightarrow{\nabla^{LC,\,A}} \Gamma(T^*W \otimes \Lambda^2_+ \mathcal{H}^* \otimes ad\,P)\\
&\xrightarrow{Cl' \otimes id} \Gamma(\mathcal{H}^* \otimes ad\,P).
\end{aligned}
$$

Here $\nabla^{LC,\,A}$ denotes the tensor product of A and the connection on $\Lambda^2_+\mathcal{H}^*$ induced by the Levi-Civita connection (we do not assume that $\Lambda^2_+\mathcal{H}^*$ is preserved by the Levi-Civita connection). It is convenient to define a map $\Phi\colon \mathcal{B}\to\Omega^1_h(ad\,P)\times\Omega^2_+(ad\,P)$ by the left hand side of Equations (3.1).

Remark 3.1. Equations (3.1) were independently discovered by Witten [30] from a different perspective. A partial case with $B\equiv 0$ has been studied by Fan [10].

Remark 3.2. The total space of $\Lambda^2_+\mathcal{H}\to W$ is an eight-manifold equipped with a natural $Spin(7)$-structure, which is induced by the Riemannian metric and orientation on W. This $Spin(7)$-structure can be constructed using the technique of [4]. Then, following the line of argument in [14], one can show that solutions of Equations (3.1) correspond to $Spin(7)$-instantons on $\Lambda^2_+\mathcal{H}$ invariant along each fibre.

The gauge group $\mathcal{G}(P)$ acts on the configuration space $\mathcal{B}$ on the right

$$(A,B)\cdot g=(A\cdot g,\ ad_{g^{-1}}B),\qquad g\in\mathcal{G}(P),$$

where g acts on the first component by the usual gauge transformation. The infinitesimal action at a point (A,B) is given by the map

$$K\colon\Omega^0(ad\,P)\longrightarrow\Omega^1(ad\,P)\oplus\Omega^2_+(ad\,P),\qquad \xi\mapsto\big(d_A\xi,\ [B,\xi]\big).$$

Notice also that the map Φ is $\mathcal{G}(P)$-equivariant.

A standard computation yields

$$\delta\Phi_{(A,\,B)}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\begin{pmatrix}\imath_v(d_A\alpha)-\delta_A^+\beta+\alpha\cdot B\\ d_A^+\alpha-\nabla_v^A\beta-[\alpha(v),B]-2\sigma(B,\beta)\end{pmatrix},\quad (\alpha,\beta)\in T_{(A,B)}\mathcal{B},$$

where the term $\alpha\cdot B\in\Omega^1_h(ad\,P)$ is constructed algebraically from α and B, namely $\alpha\cdot B=Cl'\otimes[\cdot\,,\cdot]_{\mathfrak{g}}(\alpha\otimes B)$. Thus we get the deformation complex at the point (A,B):

$$\begin{aligned}(3.2)\qquad 0\to\Omega^0(ad\,P)&\xrightarrow{K}\Omega^1(ad\,P)\oplus\Omega^2_+(ad\,P)\\ &\xrightarrow{\delta\Phi}\Omega^1_h(adP)\oplus\Omega^2_+(ad\,P)\to 0.\end{aligned}$$

Lemma 3.3. *Deformation complex (3.2) is elliptic.*

The statement of this Lemma follows immediately from Remark 3.2. Alternatively, one can consider Equations (3.1) on $\mathbb{R}^5$ and show that the symbol of $K^* + \delta\Phi$ is modelled on the octonionic multiplication. We omit the details.

4. Dimensional reductions

Before turning our attention to the dimensional reductions of Equations (3.1) a little digression is in place. Suppose a Lie group $\mathcal{G}$ acts *freely* and isometrically on a Riemannian manifold M. Identify a $\mathcal{G}$-invariant function $f\colon M \to \mathbb{R}$ with a function $\hat{f}\colon M/\mathcal{G} \to \mathbb{R}$. Then critical points of $\hat{f}$ correspond to orbits of solutions of the equation $\operatorname{grad} f = K_\xi$, where $\xi \in Lie(\mathcal{G})$ and K_ξ is the Killing vector field corresponding to ξ. But the invariance of f implies $\langle \operatorname{grad} f, K_{\xi'}\rangle = 0$ for any $\xi' \in Lie(\mathcal{G})$ so that we necessarily have $\operatorname{grad} f = 0$ for any point on M projecting to a critical point of $\hat{f}$.

Similarly, a curve $m\colon \mathbb{R} \to M$ projects to an antigradient flow of $\hat{f}$ if and only if there exists $\xi\colon \mathbb{R} \to Lie(\mathcal{G})$ such that

$$\dot{m} = -\operatorname{grad} f + K_\xi. \tag{4.1}$$

The Lie group $\{g\colon \mathbb{R} \to \mathcal{G}\}$ acts on solutions of Equation (4.1) and the orbits are in bijective correspondence with the antigradient flow lines of $\hat{f}$. Furthermore, we may consider only those solutions, which are horizontal with respect to the natural connection. This gives a bijection between ordinary flow lines of f modulo $\mathcal{G}$ and flow lines of $\hat{f}$.

The upshot is that $\mathcal{G}$-invariance of f implies that Equation (4.1) is equivalent to the ordinary antigradient flow equation of f. It will be important to switch freely between these two approaches in an infinite-dimensional setup. The reasons will be clear below.

4.1. Dimension four

Let X be a closed oriented Riemannian four-manifold. Below we consider Equations (3.1) on $(W, v) = (X \times \mathbb{R}_t, \frac{\partial}{\partial t})$ endowed with the product metric.

Denote by $pr\colon X \times \mathbb{R} \to X$ the canonical projection and set $P = pr^* P_X$, where $P_X \to X$ is a principal G-bundle. Think of $B \in \Omega^2_+(X \times \mathbb{R}; pr^* ad\, P_X)$ as a map $b\colon \mathbb{R} \to \Omega^2_+(X; ad\, P_X)$. Similarly $A \in \mathcal{A}(pr^* P_X)$ can be seen as a map $(a, c)\colon \mathbb{R} \to \mathcal{A}(P_X) \times \Omega^0(ad\, P_X)$, where c is the Higgs field. Then Equations (3.1) are easily seen to become

$$\begin{aligned} \dot{a} &= \delta_a^+ b + d_a c, \\ \dot{b} &= F_a^+ - \sigma(b, b) - [c, b], \end{aligned} \tag{4.2}$$

where $\delta_a^+ = (d_a^+)^*$. These equations turn out to be the antigradient flow equations of some function. Indeed, consider the function

$$h\colon \Lambda^2_+ H^* \otimes \mathfrak{g} \to \mathbb{R}, \qquad h(\mathrm{w}) = \frac{1}{3}\langle \mathrm{w}, \sigma(\mathrm{w})\rangle. \tag{4.3}$$

Choose an isomorphism $\Lambda^2_+ H^* \cong \mathbb{R}^3$ and write $\mathrm{w} = \sum_{i=1}^3 e_i \otimes \xi_i$. Then we have $h(\mathrm{w}) = \langle \xi_1, [\xi_2, \xi_3]\rangle$ and therefore $\operatorname{grad} h(\mathrm{w}) = \sigma(\mathrm{w})$. Since h is equivariant with respect to both $SO(3)$ and G, we obtain a well-defined map $\Omega^2_+(ad\, P_X) \to C^\infty(X)$ denoted by the same letter.

Denote $\mathcal{B} = \mathcal{A}(P) \times \Omega^2_+(ad\, P)/\mathcal{G}(P)$. As usual, $\mathcal{B}^* \subset \mathcal{B}$ denotes the quotient space of irreducible points. The negative L^2-gradient of the function

$$U\colon \mathcal{B} \to \mathbb{R}, \qquad U(a,b) = -\langle F_a^+, b\rangle_{L^2} + \int_X h(b)\, vol_X$$

is $(\delta_a^+ b,\ F_a^+ - \sigma(b,b))$. Hence, assuming there are no reducible solutions, Equations (4.2) represent the antigradient flow equations of the function $\hat{U}\colon \mathcal{B}^* \to \mathbb{R}$.

We summarize our computations in the following proposition.

Proposition 4.1. *If there are no reducible solutions, Equations (4.2) represent antigradient flow equations of the function $\hat{U}\colon \mathcal{B}^* \to \mathbb{R}$.* □

The critical points of the function U are solutions of the Vafa-Witten equations [27]:

$$\begin{aligned} &\delta_a^+ b + d_a c = 0,\\ &F_a^+ - \sigma(b,b) + [b,c] = 0. \end{aligned}$$

These equations are elliptic and the expected dimension of the moduli space is zero.

As we have seen, the $\mathcal{G}(P)$-invariance of U implies that for each irreducible solution of the Vafa-Witten equations we have $(d_a c, [b,c]) = 0$, i.e. in the absence of reducible solutions the above equations are equivalent to

$$\begin{aligned} &\delta_a^+ b = 0,\\ &F_a^+ - \sigma(b,b) = 0. \end{aligned} \tag{4.4}$$

Denote by W^+ and s the self-dual Weyl tensor and scalar curvature of X respectively. Then the Weitzenböck formula

$$2d_a^+\delta_a^+ = (\nabla^a)^*\nabla^a - 2W^+ + \frac{s}{3} + \sigma(F_a^+, \cdot),$$

yields

$$4\|\delta_a^+ b\|^2 + \|F_a^+ - \sigma(b,b)\|^2 = 2\|\nabla^a b\|^2 + \|F_a^+\|^2 + \|\sigma(b,b)\|^2 - 4\langle W^+(b), b\rangle + \frac{2}{3}\langle sb, b\rangle.$$

Proposition 4.2 ([27]). *If the operator $-W^+ + \frac{1}{6}s$ is pointwise non-negative definite on $\Lambda^2_+ T^*X$, then for any irreducible solution (a,b) of the Vafa-Witten equations the following holds: $F_a^+ = 0$, $\nabla^a b = 0$.* □

4.2. Dimension three

In this section various forms of Equations (3.1) are studied on $Y^3 \times \mathbb{R}^2$, where Y is a closed oriented Riemannian three-manifold.

Just like in the instanton Floer theory, consider solutions of (4.4) on $X = Y \times \mathbb{R}$. Assuming a is in a temporal gauge, we obtain the following system of equations

$$\begin{aligned} \dot a &= -*(F_a - \tfrac{1}{2}[b\wedge b]),\\ \dot b &= *d_a b,\\ 0 &= \delta_a b, \end{aligned} \tag{4.5}$$

where (a,b) is interpreted as a curve in $\mathcal{A}(P)\times\Omega^1(ad\,P)\cong T^*\mathcal{A}(P)$, $*$ stays for the Hodge operator on Y, and $[\cdot\wedge\cdot]$ is a combination of wedging and Lie brackets. Here we have also used the isomorphism $\Gamma(\pi^*T^*Y)\cong\Omega^2_+(Y\times\mathbb{R})$, $\omega\mapsto\frac{1}{2}(*\omega + ds\wedge\omega)$, where $\pi\colon Y\times\mathbb{R}\to Y$ is the projection.

Observe that $T^*\mathcal{A}(P)\cong\Omega^1(ad\,P)\otimes\mathbb{C}$ is a Hermitian affine space, hence a (flat) Kähler manifold. The action of the gauge group is Hamiltonian and the momentum map is given by

$$\mu\colon T^*\mathcal{A}(P)\to\Omega^0(ad\,P),\qquad \mu(a,b)=\delta_a b. \tag{4.6}$$

Denote $N=\mu^{-1}(0)=\{(a,b)\mid \delta_a b=0\}\subset T^*\mathcal{A}(P)$. It follows from the very definition of the momentum map that $d\mu$ is surjective at (a,b) if and only if the gauge group acts locally freely at (a,b). Therefore, the subset N^* consisting of all irreducible points of N is a submanifold. Hence, $N^*/\mathcal{G}(P)$ is a Kähler manifold.

Consider the map

$$f_0\colon \mathcal{A}(P)\times\Omega^1(ad\,P)\to\mathbb{R}/\mathbb{Z},\qquad f_0(a,b)=8\pi^2\vartheta(a)-\tfrac{1}{2}\langle b, *d_a b\rangle_{L^2},$$

where ϑ is the Chern-Simons function. It is easy to check that the vector field $\operatorname{grad} f_0 = \big(*(F_a - \frac{1}{2}[b\wedge b]),\ -*d_a b\big)$ is tangent to N^* at each point. Therefore critical points of the restriction of f_0 to N^* are solutions of Hitchin's equations[4] [15]:

$$F_a - \tfrac{1}{2}[b\wedge b] = 0,\quad d_a b = 0,\quad \delta_a b = 0. \tag{4.7}$$

More accurately, in the same manner as described at the beginning of this section, orbits of irreducible solutions to (4.7) correspond to critical points of $\hat f_0\colon N^*/\mathcal{G}(P)\to\mathbb{R}$. Similarly, orbits of (4.5) correspond to the flow lines of $\hat f_0$.

Remark 4.3. Denote by G^c the complexified Lie group and by $\mathcal{P} = P\times_G G^c$ the principal G^c-bundle associated with P. Any connection on $\mathcal{P}$ can be written in the form $\mathcal{A} = a + ib$, where $(a,b)\in\mathcal{A}(P)\times\Omega^1(ad\,P)$. Conversely, any pair (a,b) combines to a G^c connection $\mathcal{A}$. Then $\mathcal{A}$ is flat if and only if the first two equations of (4.7) are satisfied. The last equation, i.e. the vanishing of the moment map, has been analyzed in [6, 7].

Remark 4.4. Hitchin's equations can be obtained from $SU(3)$ anti-self-duality equations along similar lines to those outlined in Remark 3.2. Namely, the total space of T^*Y is equipped with an $SU(3)$-structure. Then $SU(3)$-instantons invariant along each fiber are solutions of Hitchin's equations.

We can also consider Equations (3.1) on $W = \mathbb{R}_t\times Y\times\mathbb{R}_s$ with $v = -\frac{\partial}{\partial s}$. Write $A = a + e\,ds + c\,dt$, where a is a family of connections on $P\to Y$. Consider first only t-invariant solutions with $c = e = 0$. A computation yields the following system:

$$\dot a = -*d_a b,\quad \dot b = -*(F_a - \tfrac{1}{2}[b\wedge b]),\quad 0 = \delta_a b, \tag{4.8}$$

where the dots denote the derivative with respect to the variable s. Equations (4.8) and (4.5) appeared in [16] for the first time and were further studied in [28, 29].

[4] Hitchin studied these equations in the case of two-dimensional base manifolds.

Consider the function

$$f_1\colon T^*\mathcal{A}(P)\to\mathbb{R},\qquad f_1(a,b)=\langle F_a,\,*b\rangle_{L^2}-\int_Y h(b)\,vol_Y,$$

where $h\colon T^*Y\otimes ad\,P\to\mathbb{R}$ is defined just like (4.3). Since grad $f_1=(*d_ab,\ *(F_a-\frac{1}{2}[b\wedge b]))$ is tangent to N^* at each point we conclude that the moduli of solutions to Equations (4.8) correspond to antigradient flow lines of $\hat{f}_1\colon N^*/\mathcal{G}(P)\to\mathbb{R}$.

Let us examine the functions f_0 and f_1 more closely. Since grad $f_1=J$grad f_0, where J is the constant complex structure on $T^*\mathcal{A}(P)\cong\Omega^1(ad\,P)\otimes\mathbb{C}$, we obtain that the function $f=f_0+if_1$ is J-holomorphic. Writing (a,b) as a G^c-connection $\mathcal{A}$ as in Remark 4.3 it is easy to check that f is the complex Chern-Simons functional

$$\mathrm{CS}(\mathcal{A})=\frac{1}{2}\int_Y\left(\langle\mathcal{A}\wedge d\mathcal{A}\rangle+\tfrac{1}{3}\langle\mathcal{A}\wedge[\mathcal{A}\wedge\mathcal{A}]\rangle\right).$$

Here we interpret $\mathcal{A}$ as a $\mathfrak{g}_{\mathbb{C}}$-valued 1-form on Y, and $\langle\cdot,\cdot\rangle\colon\mathfrak{g}_{\mathbb{C}}\otimes\mathfrak{g}_{\mathbb{C}}\to\mathbb{C}$ denotes the $\mathbb{C}$-linear extension of the scalar product on $\mathfrak{g}$.

Furthermore, let us consider Equations (3.1) on $(W,v)=(Y\times\mathbb{R}^2_{s,t},\frac{\partial}{\partial t})$. The standard reduction procedure yields the following system

$$\begin{aligned}&\partial_sa-\partial_tb+[b,c]-d_ae+*\big(F_a-\tfrac{1}{2}[b\wedge b]\big)=0,\\ &\partial_ta+\partial_sb-[b,e]-d_ac-*d_ab=0,\\ &\partial_te-\partial_sc+[c,e]+\delta_ab=0.\end{aligned}\tag{4.9}$$

Here a is a connection on the pull-back of $P=P_Y$ to $Y\times\mathbb{R}^2$, b is a 1-form and c, e are 0-forms with values in the adjoint bundle of Lie algebras. It is easy to check that these equations are symplectic vortex equations [5] with a Hamiltonian perturbation for the following data: The target space is $T^*\mathcal{A}(P)$ equipped with the Hamiltonian action of the gauge group $\mathcal{G}(P)$, $u=(a,\,b)\colon\mathbb{R}^2\to T^*\mathcal{A}(P)$, $A=e\,ds+c\,dt$, and the perturbation is $\sigma=\frac{1}{2}\mathrm{Im}\,(f dz)=\frac{1}{2}(f_0\,dt+f_1\,ds)$.

Notice also that we are free to rotate the coordinates s and t or, equivalently, to rotate the initial vector field $v=\partial_t$. This is in turn equivalent to the choice of the Hamiltonian perturbation $\sigma=\frac{1}{2}\mathrm{Im}\,(e^{i\theta}f dz)$ and the resulting equations are

$$\begin{aligned}&\partial_sa-\partial_tb+[b,c]-d_ae-\sin\theta\,*d_ab+\cos\theta\,*\big(F_a-\tfrac{1}{2}[b\wedge b]\big)=0,\\ &\partial_ta+\partial_sb-[b,e]-d_ac-\cos\theta\,*d_ab-\sin\theta\,*\big(F_a-\tfrac{1}{2}[b\wedge b]\big)=0,\\ &\partial_te-\partial_sc+[c,e]+\delta_ab=0.\end{aligned}\tag{4.10}$$

Remark 4.5. The above description of Equations (4.10) is analogous to the interpretation of the anti-self-duality equations on $\mathbb{C}\times\Sigma$ as symplectic vortex equations [5]. A new phenomenon here is the appearance of the Hamiltonian perturbation. Notice also that the adiabatic limit procedure as in [5, 13] for Equations (4.10) yields (at least formally) holomorphic planes to $T^*\mathcal{A}(P)/\!\!/\mathcal{G}(P)$ with a Hamiltonian perturbation.

For solutions of Equations (4.10) invariant with respect to s we obtain the following system

$$\begin{aligned} \dot a &= \cos\theta\, *d_a b + \sin\theta\, *\big(F_a - \tfrac{1}{2}[b\wedge b]\big) + d_a c + [b,e],\\ \dot b &= -\sin\theta\, *d_a b + \cos\theta\, *\big(F_a - \tfrac{1}{2}[b\wedge b]\big) - d_a e + [b,c],\\ \dot e &= -\delta_a b + [e,c]. \end{aligned} \tag{4.11}$$

Notice that if $c=e=0$ we obtain Equations (4.5) and (4.8) for $\theta=-\pi/2$ and $\theta=\pi$, respectively.

It will be helpful in the sequel to consider Equations (4.11) from a more abstract point of view. Namely, let (M,ω) be a symplectic manifold. Assume a Lie group $\mathcal{G}$ acts on M in a Hamiltonian manner. Denote by $\mu\colon M\to\mathfrak{G}=Lie(\mathcal{G})$ the corresponding moment map. Let $f\colon M\to\mathbb{C}$ be a $\mathcal{G}$-invariant J-holomorphic function, where J is a $\mathcal{G}$-invariant almost complex structure on M. Consider the following equations for a curve (γ,ξ,η) in $M\times\mathfrak{G}\times\mathfrak{G}$:

$$\begin{aligned} \dot\gamma &= \sin\theta\,\mathrm{grad}\, f_0(\gamma) + \cos\theta\,\mathrm{grad}\, f_1(\gamma) + K_\xi(\gamma) - JK_\eta(\gamma),\\ \dot\eta &= -\mu(\gamma) - [\xi,\eta], \end{aligned} \tag{4.12}$$

where K is the Killing vector field. Clearly, we obtain Equations (4.11) from (4.12) putting $M=T^*\mathcal{A}(P)$, $\gamma=(a,b), \eta=e$, and $\xi=c$.

Notice that as explained in the beginning of Section 4 the $\mathcal{G}$-invariance of f implies that for any $\xi\in\mathfrak{G}$ we have $\langle \mathrm{grad}\, f_i, K_\xi\rangle = 0,\ i=0,1$. Then, using the definition of the moment map, for any solution of (4.12) we obtain

$$\begin{aligned} \frac{d}{dt}\langle\mu(\gamma),\eta\rangle &= \omega(K_\eta,\dot\gamma) + \langle\mu,\dot\eta\rangle\\ &= g(JK_\eta,K_\xi) - g(JK_\eta,JK_\eta) - \langle\mu,\mu\rangle - \langle\mu,[\xi,\eta]\rangle. \end{aligned} \tag{4.13}$$

Furthermore, observe that for any $\zeta,\rho\in\mathfrak{G}$ the following equalities hold:

$$\begin{aligned} d\big(\omega(K_\zeta,K_\rho)\big) &= d\,\imath_{K_\rho}(\imath_{K_\zeta}\omega) = \mathcal{L}_{K_\rho}(\imath_{K_\zeta}\omega) - \imath_{K_\rho} d(\imath_{K_\zeta}\omega) = \imath_{K_{[\rho,\zeta]}}\omega\\ &= -d\,\langle\mu,[\zeta,\rho]\rangle. \end{aligned}$$

Here the second equality follows from Cartan's equation. Hence, $\langle \mu, [\zeta, \rho] \rangle = -\omega(K_\zeta, K_\rho) = g(K_\zeta, JK_\rho)$. Therefore, by (4.13) we have

$$\frac{d}{dt}\langle \mu(\gamma), \eta \rangle = -g(K_\eta, K_\eta) - \langle \mu, \mu \rangle \leq 0. \tag{4.14}$$

Hence, for any solution of Equations (4.12) the function $\langle \mu(\gamma), \eta \rangle$ is non-increasing.

We will be interested below in solutions (γ, ξ, η) of (4.12) satisfying the condition

$$(\gamma, \xi, \eta) \longrightarrow (m_\pm, 0, 0) \qquad \text{as} \quad t \to \mp\infty, \tag{4.15}$$

where $m_\pm$ are critical points of f. For any such solution $\langle \mu(\gamma), \eta \rangle$ vanishes at $\pm\infty$ and hence vanishes everywhere. Then from (4.14) we conclude that η and $\mu \circ \gamma$ vanish everywhere, i.e. under condition (4.15) Equations (4.12) reduce to

$$\dot{\gamma} = \sin\theta \operatorname{grad} f_0 + \cos\theta \operatorname{grad} f_1 + K_\xi, \qquad \mu(\gamma) = 0.$$

From the discussion at the beginning of Section 4 we obtain that these equations are equivalent to

$$\dot{\gamma} = \sin\theta \operatorname{grad} f_0 + \cos\theta \operatorname{grad} f_1, \qquad \mu(\gamma) = 0. \tag{4.16}$$

Summing up, we have that under condition (4.15) systems (4.12) and (4.16) are equivalent. Applying this conclusion in the case $M = T^*\mathcal{A}(P)$ we obtain that for $\theta = -\pi/2$ Equations (4.11) together with the condition

$$(a, b, c, e) \longrightarrow (a_\pm, b_\pm, 0, 0) \qquad \text{as} \quad t \to \pm\infty,$$

where $(a_\pm, b_\pm)$ are solutions of Hitchin's equations, are equivalent to Equations (4.5) together with $(a, b) \to (a_\pm, b_\pm)$ as $t \to \mp\infty$. The upshot is that while Equations (4.5) and (4.11) with $\theta = -\pi/2$ are essentially equivalent, only the latter are elliptic (this is similar to the fact that on compact three-manifolds the equation $F_A = 0$ (flat $SU(2)$-connections) is essentially equivalent to the Bogomolny equations, however only the Bogomolny equations are elliptic).

Remark 4.6. One obtains an elliptic form of Hitchin's equations on a three manifold by considering solutions of Equations (3.1) on $(Y \times \mathbb{R}^2, \partial_t)$ invariant along $\mathbb{R}^2$. The corresponding equations are easily obtained from (4.9).

5. Invariants

In this section we outline constructions of invariants assigned to five-, four-, and three- manifolds arising from gauge theories described in the preceding sections. It is clear that the constructions described below need an appropriate analytic justification. We postpone this to subsequent papers and restrict ourselves to some examples. Throughout this section the coefficient ring is $\mathbb{Z}/2\mathbb{Z}$ in all constructions for the sake of simplicity.

The expected dimension of the moduli space of solutions of Eqs. (3.1) for closed five-manifolds is zero. Therefore, assuming compactness and transversality, an algebraic count associates a number to closed five-manifolds. More accurately, this number depends on the isomorphism class of P and on the class of the vector field v in $\pi_0(\mathfrak{X}_0(W))$, where $\mathfrak{X}_0(W)$ denotes the space of all vector fields on W without zeros.

Let us now consider the dimension four. The corresponding construction is very similar to the instanton Floer theory, so we are very brief here. Assume the moduli space of solutions to the Vafa-Witten equations $\mathcal{M}_{VW}$ is compact and zero-dimensional (for the case $\dim \mathcal{M}_{VW} > 0$ see example below). The index of the Hessian of U on $X^4 \times S^1$ vanishes and therefore the relative Morse index of a pair of critical points is an integer.[5] The Floer differential counts the moduli space of finite-energy solutions of Equations (3.1) on $X \times \mathbb{R}$ converging to solutions of the Vafa-Witten equations at $\pm\infty$. As a result, for a smooth four-manifold equipped with a principal G-bundle Floer-type homology groups can conjecturally be constructed.

Example 5.1. Let X be a Kähler surface with a non-negative scalar curvature. Then Proposition 4.2 applies and, therefore, $\mathcal{M}_{VW} = \mathcal{M}_{asd}$ assuming all asd connections are irreducible and non-degenerate. If $\dim \mathcal{M}_{asd} > 0$ the function U is not Morse but rather Morse-Bott. Then choosing a suitable perturbation, which is essentially a Morse function φ on $\mathcal{M}_{asd}$, one obtains the Morse-Witten complex of φ. The details can be found for instance in [2]. In other words, the corresponding Floer homology groups are homology groups of $\mathcal{M}_{asd}$. Notice that this agrees perfectly with the Vafa-Witten theory: The Vafa-Witten invariant, which counts solutions of the Vafa-Witten equations, is the Euler characteristic of $\mathcal{M}_{asd}$ provided the only solutions of the Vafa-Witten equations are anti-self-dual instantons.

[5] In general, there is no a distinguished critical point as in the $SU(2)$-instanton Floer theory, so that we are left with the relative grading only.

It is worth pointing out that the above reasoning is valid if $\mathcal{M}_{asd}$ admits a compactification, which is a manifold. Notice that the Euler characteristic of $\mathcal{M}_{asd}$ in [27] is taken as the Euler characteristic of the Gieseker compactification.

Furthermore, let us consider dimension three. Let (Y, g) be a closed oriented Riemannian three-manifold. Pick a nontrivial principal G-bundle $P \to Y$ and assume that all solutions of Hitchin's equations are irreducible (thus we exclude the case $G = SU(2)$) and the moduli space is finite, say $\{\mathcal{A}_1, \dots, \mathcal{A}_k\}$. Recall that this is the critical set of the complex Chern-Simons functional and therefore we can conjecturally construct a corresponding collection of k Fukaya-Seidel A_∞-categories[6] $\mathcal{A}_j(Y)$ as described in Section 2.2.

Thus, the objects of $\mathcal{A}_j(Y)$ are classes of solutions $\mathcal{A}_l$ of Hitchin's equations. For ease of exposition we assume that $\operatorname{Re}\operatorname{CS}(\mathcal{A}_1) < \cdots < \operatorname{Re}\operatorname{CS}(\mathcal{A}_k)$, where $\operatorname{Re}\operatorname{CS}(\mathcal{A}_l)$ is understood to take values in $[0, 1)$. Recall that for any pair $\mathcal{A}_\pm \in \{\mathcal{A}_1, \dots, \mathcal{A}_k\}$, $\mathcal{A}_- < \mathcal{A}_+$, the space $hom(\mathcal{A}_-, \mathcal{A}_+)$ is generated by the broken flow lines of the complex Chern-Simons functional connecting $\mathcal{A}_-$ with $\mathcal{A}_+$. More precisely, as described in Remark 2.5, we consider only those broken flow lines γ for which the image of $\operatorname{CS} \circ \gamma$ does not intersect the set $\big(\operatorname{Re}\operatorname{CS}(\mathcal{A}_j), \operatorname{Re}\operatorname{CS}(\mathcal{A}_{j+1})\big) \times \mathbb{R}$. Recall also that the flow lines of the complex Chern-Simons functional can conveniently be described as moduli of solutions of Equations (4.11) satisfying the asymptotic conditions

$$(a, b, c, e) \longrightarrow (a^0_\pm, b^0_\pm, 0, 0) \qquad \text{as} \quad t \to \mp\infty, \tag{5.1}$$

where $(a^0_\pm, b^0_\pm)$ are solutions of Hitchin's equations representing $\mathcal{A}_\pm$.

Furthermore, the Floer differential $\mu^1 \colon hom(\mathcal{A}_-, \mathcal{A}_+) \to hom(\mathcal{A}_-, \mathcal{A}_+)$ is obtained by counting moduli of finite-energy pseudoholomorphic planes with a Hamiltonian perturbation satisfying suitable conditions at infinity. In our case, by Remark 4.5 these pseudoholomorphic planes can (formally) be interpreted as solutions of Equations (4.10), which are in turn interpreted as solutions of Equations (3.1) on $W = Y \times \mathbb{R}^2$.

Summing up, choose any admissible pair $\mathcal{B}_\pm$ of gauge equivalence classes of finite-energy solutions of Equations (4.11) and (5.1). Then *define* the map μ^1 by counting moduli of solutions to Equations (3.1) on $(W, v) = (Y \times \mathbb{R}^2, \cos\theta\,\partial_t + \sin\theta\,\partial_s)$ with the following boundary conditions

$$\begin{aligned} (a, b, c, e) &\to (a_\pm(t), b_\pm(t), 0, 0) \qquad && \text{as} \quad s \to \mp\infty, \\ (a, b, c, e) &\to (a^0_\pm, b^0_\pm, 0, 0) && \text{as} \quad t \to \mp\infty, \end{aligned}$$

[6] $\mathcal{A}_j(Y)$ will also depend on the metric as well as on the choice of P.

where $(a_\pm(t), b_\pm(t))$ represents the class $\mathcal{B}_\pm$.

To define the map μ^2, one considers finite-energy solutions of Equations (3.1) on $W = Y \times \Omega$ satisfying appropriate boundary conditions, where Ω is as shown in Fig.1. The maps μ^d for $d \geq 3$ are defined similarly and conjecturally the whole collection $\{\mu^d\}$ combines to an A_∞-structure.

Notice that the change of orientation on Y is equivalent to multiplication of f by -1 and hence does not affect $\mathcal{A}_j(Y)$. On the other hand, $\mathcal{A}_j(Y)$ depends on the Riemannian metric g. However, as explained in [24] the derived category $D^b(\mathcal{A}_j(Y))$ should be independent of g.

Appendix A. Pseudoholomorphic strips and pseudoholomorphic planes

In this appendix we outline (without proof) a connection between pseudoholomorphic planes with a Hamiltonian perturbation and pseudoholomorphic strips with Lagrangian boundary conditions. To do so, pick a pair (m_-, m_+) of critical points of f and assume that the interval $\overline{z_- z_+}$ does not contain any other critical point, where $z_\pm = f(m_\pm)$. It is convenient to choose the midpoint of $\overline{z_- z_+}$ as the basepoint. We deviate here from our convention on the choice of the basepoint for the convenience of exposition only, namely to avoid differential equations with non-smooth coefficients.

Replacing f with $e^{-i\theta_\pm}(f - z_0)$ if necessary we may assume that $z_\pm = \pm T, T > 0$ and hence $z_0 = 0, \theta_0(t) \equiv 0$. We establish a relation between solutions of the equations

$$\text{(A.1)} \qquad \begin{aligned} &\partial_s u + J(\partial_t u + v_0) = 0, \qquad u\colon \mathbb{R}^2_{s,t} \to M,\\ &\lim_{t\to\pm\infty} u(s,t) = m_\mp, \qquad \lim_{s\to\pm\infty} u(s,t) = \gamma_\mp(t) \end{aligned}$$

and pseudoholomorphic strips in two steps. In the first step we relate solutions of Equations (A.1) to solutions of the problem

$$\text{(A.2)} \qquad \begin{aligned} &\partial_s u_0 + J\Big(\partial_\tau u_0 + \frac{1}{\|v_0\|^2} v_0\Big) = 0, \qquad (s,\tau) \in \mathbb{R} \times (-T, T),\\ &u_0(s, \pm T) = m_\mp, \qquad \lim_{s\to\pm\infty} u_0(s,\tau) = \gamma_{0,\pm}(\tau), \end{aligned}$$

where $\gamma_{0,\pm}$ satisfies the equations

$$\text{(A.3)} \qquad \begin{aligned} &\frac{d}{d\tau}\gamma_0 + \frac{1}{\|v_0\|^2} v_0 = 0, \qquad \tau \in (-T, T),\\ &\gamma_0(\pm T) = m_\mp. \end{aligned}$$

In the second step we show how to relate solutions of (A.2) to pseudoholomorphic strips.

Step 1. It is an elementary fact that Equations (A.3) are equivalent to the antigradient flow equations for f_0. Nevertheless it is instructive to examine this equivalence more closely. Consider the family of equations

$$\text{(A.4)} \qquad \begin{aligned} &\dot\gamma_\lambda + \frac{1}{\lambda + (1-\lambda)\|v_0\|^2}\, v_0 = 0, \qquad \gamma_\lambda \colon \mathbb{R} \to M, \\ &\lim_{t\to\pm\infty} \gamma_\lambda(t) = m_\mp, \end{aligned}$$

where $\lambda \in (0,1]$, and fix a parametrization by the condition $f_0 \circ \gamma_\lambda(0) = 0$. Pick any solution γ_1 of Equations (A.4) for $\lambda = 1$, i.e. an antigradient flow line of f_0, and consider the following family of diffeomorphisms

$$\tau_\lambda \colon \mathbb{R} \to \mathbb{R}, \quad \tau_\lambda(t) = \lambda t + (1-\lambda) f_0 \circ \gamma_1(t), \qquad \lambda \in (0,1].$$

It is straightforward to check that $\gamma_\lambda = \gamma_1 \circ \tau_\lambda^{-1}$ is a solution of (A.4), and this establishes a bijection between antigradient flow lines of f_0 and solutions of (A.4). This correspondence is also valid for $\lambda = 0$, but in this case τ_0 maps $\mathbb{R}$ bijectively onto the interval $(-T, T)$. If we extend γ_0 by the constant values outside $(-T,T)$, then γ_λ converges to γ_0 in $C^0(\mathbb{R}; M)$ as $\lambda \to 0$ (in fact, in any reasonable topology).

With this understood, consider the family of equations

$$\text{(A.5)} \qquad \begin{aligned} &\partial_s u_\lambda + J\Big(\partial_t u_\lambda + \frac{1}{\lambda + (1-\lambda)\|v_0\|^2}\, v_0\Big) = 0, \qquad (s,t) \in \mathbb{R}^2 \\ &\lim_{t\to\pm\infty} u_\lambda(s,t) = m_\mp, \qquad \lim_{s\to\pm\infty} u_\lambda(s,t) = \gamma_{\lambda,\pm}(t). \end{aligned}$$

For these equations explicit correspondence between solutions for different values of λ is not available anymore, but it is reasonable to expect that u_λ converges to a solution of (A.2) as $\lambda \to 0$.

Step 2. Let $L_\pm(\tau) \subset f^{-1}(\tau),\ \tau \in (-T,T)$, denote the vanishing cycle of $m_\pm$ associated with the segment $[\tau, \pm T]$. Consider the family of equations

$$\begin{aligned} &\partial_s u_\mu + J\Big(\partial_t u_\mu + \frac{1-\mu}{\|v_0\|^2}\, v_0\Big) = 0, \qquad (s,\tau) \in \mathbb{R} \times (-T,T) \\ &u_\mu(s, \pm T) \in L_\pm(\pm(1-\mu)T), \qquad \lim_{s\to\pm\infty} u_\mu(s,\tau) = \gamma_{0,\pm}((1-\mu)\tau) \end{aligned}$$

with $\mu \in [0,1]$. Clearly, for $\mu = 0$ we obtain Equations (A.2), whereas for $\mu = 1$ we have holomorphic strips as in the classical definition of the Floer differential. Notice that the images of such holomorphic strips lie in the fiber of f. This follows from the J-holomorphic property of f and the maximum principle (see also Remark A.1 for more details).

Remark A.1. Pick a solution u_0 of Equations (A.2) and denote $f\circ u_0 = \varphi + i\psi$. It follows from the holomorphicity of f that φ and ψ satisfy the inhomogeneous Cauchy-Riemann equations

$$\partial_s\varphi - \partial_\tau\psi = 0, \quad \partial_s\psi + \partial_\tau\varphi + 1 = 0$$

and therefore both functions are harmonic. Moreover, the holomorphicity of f also implies that $f_1\circ\gamma_{0,\pm}(\tau)$ is constant it τ. Since $f_1\circ\gamma_\pm(\pm T) = f_1(m_\pm) = 0$, $f_1\circ\gamma_{0,\pm}$ vanishes everywhere. We conclude that ψ vanishes as $\tau\to\pm T$ and as $s\to\pm\infty$ and thus vanishes everywhere. Therefore $\varphi(s,\tau) = -\tau$. We see that unlike pseudoholomorphic strips, images of solutions of (A.2) do not lie in a fixed fiber of f, but rather we have $u_0(s,\tau)\in f^{-1}(-\tau)$.

Notice also that at the first glance Equation (A.2) has singularities. Namely, if a solution u_0 hits a critical point of f at a single point (s_0,τ_0), then φ and ψ are harmonic in $\mathbb{R}\times(-T,T)\setminus\{(s_0,\tau_0)\}$ and continuous at (s_0,τ_0). Hence the singularity is removable and the above argument shows that the image of $f\circ u_0$ is the segment $(-T,T)$. Since by assumption the segment $(-T,T)$ does not contain any critical values, we conclude that a priori a solution of (A.2) cannot hit a critical point of f in an interior point.

Appendix B. On broken flow lines

In this appendix missing details on broken flow lines are provided. We use notations introduced in Subsections 2.2 and 2.3.

Lemma B.1. *Suppose the closed domain G bounded by the triangle $z_-z_0z_+$ contains no critical values of f other than $z_\pm$. Denote by ℓ the curve $\overline{z_-z_0}\cup\overline{z_0z_+}$. Then for any $\varepsilon>0$ there exists $\nu_0>0$ such that for all broken flow lines γ_ν of f connecting m_- and m_+ and all $t\in\mathbb{R}$ we have*

$$d\big(f\circ\gamma_\nu(t),\,\ell\big)<\varepsilon \qquad \textit{provided} \quad \nu\le\nu_0. \tag{B.1}$$

Proof. The lemma is proved in three steps.

Step 1. For any broken flow line γ_ν the image of the curve $f\circ\gamma_\nu\colon\mathbb{R}\to\mathbb{C}$ is contained in G.

From (2.6) we have $\frac{d}{dt}\operatorname{Im} f\circ\gamma_\nu(t) = -\sin\theta_\nu(t)\rho\circ\gamma_\nu\le 0$ for $t\le 0$. Since $\lim_{t\to-\infty} f\circ\gamma_\nu(t) = z_+$ we conclude that $\operatorname{Im} f\circ\gamma_\nu(t)\le\operatorname{Im} z_+ = \zeta$ for all

$t \le 0$. Similarly, recalling (2.13) we obtain $\operatorname{Im} f \circ \gamma_\nu(t) \le \operatorname{Im} z_- = \zeta$ for all $t \ge 0$. Hence, the image of the curve $f \circ \gamma_\nu$ lies in the half-plane, which is bounded by the straight line through z_- and z_+ and contains z_0. Arguing along similar lines, one also obtains that the image of $f \circ \gamma_\nu$ is contained in the half-plane bounded by the straight line through $z_\pm$ and z_0 and containing $z_\mp$.

Step 2. For any $\varepsilon > 0$ there exist $T_\varepsilon > 0$ and $\nu_0 = \nu_0(\varepsilon) > 0$ such that $f \circ \gamma_\nu(\pm t) \in B_\varepsilon(z_\mp)$ for all $t \ge T_\varepsilon$ and all $\nu \le \nu_0$.

We prove that $f \circ \gamma_\nu(t) \in B_\varepsilon(z_-)$ for all $t \ge T_\varepsilon$ and all $\nu \le \nu_0$. The rest can be proved similarly.

For an arbitrary $\varepsilon > 0$ denote

$$\begin{aligned} &\rho_\varepsilon = \inf\bigl\{\rho(m) \mid f(m) \in G,\ \operatorname{Im} f(m) \le \zeta - \varepsilon\bigr\} > 0,\\ &T_\varepsilon = 1 + \frac{1+\zeta}{\rho_\varepsilon \sin\theta_-}, \qquad \nu_0 = T_\varepsilon^{-1}. \end{aligned}$$

We claim that $\operatorname{Im} f \circ \gamma_\nu(T_\varepsilon) > \zeta - \varepsilon$ for any $\nu \le \nu_0$. Indeed, assume this is not the case, i.e. there exists $\nu \le \nu_0$ such that $\operatorname{Im} f \circ \gamma_\nu(T_\varepsilon) \le \zeta - \varepsilon$. Then for any $t \in [1, T_\varepsilon]$ we have $\operatorname{Im} f \circ \gamma_\nu(t) \le \zeta - \varepsilon$ since the function $\operatorname{Im} f \circ \gamma_\nu(t)$ is monotone for $t \ge 0$ as indicated in the proof of Step 1. Hence,

$$\begin{aligned} \zeta &\ge \operatorname{Im}\bigl(f \circ \gamma_\nu(T_\varepsilon) - f \circ \gamma_\nu(1)\bigr) = \sin\theta_- \int_1^{T_\varepsilon} \rho \circ \gamma_\nu(t)\, dt \\ &\ge \sin\theta_- \rho_\varepsilon (T_\varepsilon - 1) \ge 1 + \zeta. \end{aligned}$$

This contradiction proves the inequality $\operatorname{Im} f \circ \gamma_\nu(T_\varepsilon) > \zeta - \varepsilon$, which in turn implies that $\operatorname{Im} f \circ \gamma_\nu(t) > \zeta - \varepsilon$ for all $t \ge T_\varepsilon$ and all $\nu \le \nu_0$. Arguing along similar lines and redenoting T_ε, ν_0 if necessary one also obtains that the inequality $\operatorname{Im}\bigl(e^{-i\theta_+} f \circ \gamma_\nu(t)\bigr) \ge \operatorname{Re}(e^{-i\theta_+} z_-) - \varepsilon$ holds for all $t \ge T_\varepsilon$ and all $\nu \le \nu_0$. This implies Step 2.

Step 3. We prove the lemma.

Pick any $\varepsilon > 0$. Then by Step 2 there exist $T_\varepsilon > 0$ and $\nu_0 \le T_\varepsilon^{-1}$ such that $\bigl| f \circ \gamma_\nu(t) - z_- \bigr| < \varepsilon$ holds for all $t \ge T_\varepsilon$ provided $\nu \le \nu_0$. Since for $t \in [\nu, \nu^{-1}]$ we have that $f \circ \gamma_\nu(t)$ lies on a straight line parallel to the straight line through z_0 and z_-, we obtain that inequality (B.1) holds for all $t \ge \nu$. Using similar arguments one shows that inequality (B.1) also holds for $t \le -\nu$. Furthermore, the length of the curve $f \circ \gamma_\nu(t),\ t \in [-\nu, \nu]$, is bounded by $2\sqrt{\bar\rho}\nu$, where $\bar\rho = \sup\{\rho(m) \mid m \in M\}$. Redenoting ν_0 if necessary we obtain that inequality (B.1) holds for $t \in [-\nu, \nu]$ as well. □

Proof of Proposition 2.6. The proof consists of the following four steps

Step 1. For any $\varepsilon > 0$ there exists $\nu_0 > 0$ such that $d(\gamma_\nu(t), m_-) < \varepsilon$ for all $\nu \le \nu_0,\ t \ge \nu^{-1} + 1$, and $\gamma_\nu \in \Gamma_\nu(m_-, m_+)$.

From the equality $\frac{d}{dt}\mathrm{Re}\,(f \circ \gamma_\nu(t)) = -|\dot\gamma_\nu(t)|^2$, which is valid for all $t \ge \nu^{-1} + 1$, we obtain

$$\int_{\nu^{-1}+1}^{\infty} |\dot\gamma_\nu(t)|^2\, dt = \mathrm{Re}\,(f \circ \gamma_\nu(\nu^{-1} + 1)) - \mathrm{Re}\, z_- \tag{B.2}$$

Hence, the map $\beta_\nu(t) = \gamma_\nu(\nu^{-1} + 1 + t)$ belongs to $W^{1,2}(\mathbb{R}_+; M)$. Moreover, by Step 2 in the proof of Lemma B.1 there exists $\nu_0 = \nu_0(\varepsilon)$ such that $\|\beta_\nu\|_{W^{1,2}} < \varepsilon$ for all $\nu \le \nu_0$. Hence, Step 1 follows from the Sobolev embedding $W^{1,2}(\mathbb{R}_+; M) \hookrightarrow C^0(\mathbb{R}_+; M)$.

Step 2. For any $\varepsilon > 0$ there exists $\nu_0 = \nu_0(\varepsilon) > 0$ such that $|\dot\gamma_\nu(t)| < \varepsilon$ for all $\nu \le \nu_0,\ t \in [\nu^{-1}, \nu^{-1} + 1]$, and $\gamma_\nu \in \Gamma_\nu(m_-, m_+)$.

By choosing local coordinates we can identify a neighbourhood of m_- with $\mathbb{R}^{2n}$. Since m_- is a nondegenerate critical point of $f_0 = \mathrm{Re}\, f$, we can assume that f_0 is a quadratic function in the local representation. Hence there exist positive constants C and δ such that the inequality $\rho(x) \le C^2|x|^2$ holds whenever $|x| \le \delta$. Here $|\cdot|$ is the standard Euclidean norm on $\mathbb{R}^{2n}$. Hence,

$$|\dot\gamma_\nu|^2 = \rho(\gamma_\nu) \le C^2|\gamma_\nu|^2 \tag{B.3}$$

provided $|\gamma_\nu| \le \delta$. Therefore, for any $\tau \ge 0$ we have

$$\begin{aligned} |\gamma_\nu(\nu^{-1} + \tau)| &= \Big|\gamma_\nu(\nu^{-1} + 1) - \int_{\nu^{-1}+\tau}^{\nu^{-1}+1} \dot\gamma_\nu(t)\, dt\Big| \\ &\le |\gamma_\nu(\nu^{-1} + 1)| + C \int_{\nu^{-1}+\tau}^{\nu^{-1}+1} |\gamma_\nu(t)|\, dt \end{aligned} \tag{B.4}$$

provided $|\gamma_\nu(t)| < \delta$ for all $t \in [\nu^{-1} + \tau, \nu^{-1} + 1]$.

For any $\varepsilon > 0$ such that $\varepsilon/C < \delta$ by Step 1 we can choose $\nu_0 > 0$ so small that the inequality

$$|\gamma_\nu(\nu^{-1} + 1)| < \frac{e^{-C}\varepsilon}{C} < \delta$$

holds for all $\nu \le \nu_0$ and $\gamma_\nu \in \Gamma_\nu(m_-, m_+)$. In particular, (B.4) holds for all τ sufficiently close to 1. By the Gronwall-Bellman inequality we obtain

$$|\gamma_\nu(\nu^{-1} + \tau)| \le |\gamma_\nu(\nu^{-1} + \tau)| e^{C(1-\tau)} < \varepsilon/C \tag{B.5}$$

for all $\tau \in [0,1]$ sufficiently close to 1. This implies in fact that (B.5) holds for all $\tau \in [0,1]$. Then Step 2 follows by (B.3).

Step 3. For any $\varepsilon > 0$ *there exists* $T_\varepsilon > 0$ *and* $\nu_0 \le T_\varepsilon^{-1}$ *such that* $d(\gamma_\nu(t), m_-) < \varepsilon$ *for all* $\nu \le \nu_0,\ t \ge T_\varepsilon,$ *and* $\gamma_\nu \in \Gamma_\nu(m_-, m_+)$.

Let T_ε be as in Step 2 in the proof of Lemma B.1. From the equality $\frac{d}{dt}\mathrm{Re}\left(e^{-i\theta_-} f\circ\gamma_\nu(t)\right) = |\dot\gamma_\nu(t)|^2$, which is valid for all $t \in [\nu, \nu^{-1}]$, we obtain

$$\int_{T_\varepsilon}^{\nu^{-1}} |\dot\gamma_\nu(t)|^2\, dt = \left|\mathrm{Re}\left(e^{i\theta_-} f\circ\gamma_\nu(\nu^{-1})\right) - \mathrm{Re}\left(e^{i\theta_-} f\circ\gamma_\nu(T_\varepsilon)\right)\right| \le 2\varepsilon$$

By (B.2) and Step 2 we obtain

$$\begin{aligned}\int_{T_\varepsilon}^{\infty} |\dot\gamma_\nu|^2\, dt &= \int_{T_\varepsilon}^{\nu^{-1}} |\dot\gamma_\nu|^2\, dt + \int_{\nu^{-1}}^{\nu^{-1}+1} |\dot\gamma_\nu|^2\, dt + \int_{\nu^{-1}+1}^{\infty} |\dot\gamma_\nu|^2\, dt\\ &\le 2\varepsilon + \varepsilon^2 + \varepsilon.\end{aligned}$$

Step 3 follows from the embedding $W^{1,2}(\mathbb{R}_+; M) \hookrightarrow C^0(\mathbb{R}_+; M)$.

Step 4. We prove the proposition.

Since ρ is bounded on M, we obtain that $|\dot\gamma_j|^2 \le \sup\rho < \infty$. Then with the help of the Ascoli-Arzela theorem we can find a subsequence γ_{j_k}, which converges to some $\gamma_0 \in C^0(\mathbb{R}; M)$ on each compact interval. Then $\gamma_0 \in C^1(\mathbb{R}\setminus\{0\}; M)$ and satisfies $\dot\gamma_0 + \cos\theta_0\, v_0 + \sin\theta_0\, v_1 = 0$. By Lemma B.1 the image of $f\circ\gamma_0$ is contained in $\ell = z_- z_0 z_+$. Hence, $\lim_{t\to\pm\infty}\gamma_0(t)$ must be a critical point of the vector field $\cos\theta_\mp v_0 + \sin\theta_\mp v_1$. Hence, $\lim_{t\to\pm\infty}\gamma_0(t) = m_\mp$, i.e. γ_0 is a solution of (2.5).

Choose any $\varepsilon > 0$ sufficiently small. By Step 3 there exist $T_\varepsilon > 0$ and $N_\varepsilon > 0$ such that

$$d(\gamma_{j_k}(t), \gamma_0(t)) < \varepsilon \tag{B.6}$$

provided $t \ge T_\varepsilon$ and $k \ge N_\varepsilon$. Arguing similarly, we also obtain that (B.6) holds for $t \le -T_\varepsilon$ (possibly after increasing T_ε). Since γ_{j_k} converges on $[-T_\varepsilon, T_\varepsilon]$ we can find $N'_\varepsilon \ge N_\varepsilon$ such that (B.6) also holds for $t \in [-T_\varepsilon, T_\varepsilon]$ provided $j_k \ge N'_\varepsilon$. This finishes the proof. □

Introduce the Banach manifold

$$W^{1,2}_{m_-;\,m_+} = \{\,\gamma \in W^{1,2}(\mathbb{R}; M) \mid \lim_{t\to\pm\infty}\gamma(t) = m_\mp\}$$

and the vector bundle $\mathcal{E} \to W^{1,2}_{m_-;\,m_+}$, whose fiber at γ is the Hilbert space $L^2(\gamma^* TM)$. The map σ_ν given by (2.24) can be interpreted as a section of $\mathcal{E}$. Similarly, the map (2.25) can be interpreted as the covariant derivative of σ_ν. Then σ_ν is a Fredholm section [22] with vanishing index, since the Morse indices of m_+ and m_- are equal. Here $m_\pm$ is regarded as a critical point of $\operatorname{Re} f$. Clearly, $\sigma_\nu^{-1}(0) = \Gamma_\nu(m_-; m_+)$.

Lemma B.2. *Let $L_\pm$ be the vanishing cycle corresponding to the segment $\overline{z_\pm z_0}$. If L_+ and L_- intersect transversely in $f^{-1}(0)$, then there exists $\nu_0 > 0$ such that σ_ν intersects the zero section transversely for all $\nu \in (0, \nu_0)$. Moreover there exists a natural bijective correspondence between $\Gamma_\nu(m_-; m_+)$ and $\Gamma_0(m_-; m_+)$ provided $\nu \in (0, \nu_0)$.*

Proof. Let $\mathcal{U}_+$ denote the unstable manifold of m_+ regarded as a critical point of $\operatorname{Re}(e^{-i\theta_+} f)$. Similarly, let $\mathcal{S}_-$ denote the stable manifold of m_- regarded as a critical point of $\operatorname{Re}(e^{-i\theta_-} f)$.

Pick a point $m \in L_- \cap L_+ \cong \mathcal{S}_- \cap \mathcal{U}_+$ and observe that $\mathcal{S}_-$ and $\mathcal{U}_+$ are the Lagrangian thimbles of m_- and m_+ associated with the segments $\overline{z_0 z_-}$ and $\overline{z_0 z_+}$, respectively. Here z_0 is the origin. Then the hypothesis of the lemma implies that $\mathcal{S}_-$ and $\mathcal{U}_+$ intersect transversally at m.

Let γ_0 be the solution of (2.5) corresponding to m. Denote by $D_{\gamma_0}\sigma_0$ the linearization of σ_0 at γ_0. As we have already remarked above, $D_{\gamma_0}\sigma_0 : W^{1,2}(\gamma_0^* TM) \to L^2(\gamma_0^* TM)$ is a Fredholm operator of index 0. Moreover, it can be shown in the similar manner as in the proof of Theorem 3.3 in [22] that $\dim \operatorname{coker} D_{\gamma_0}\sigma_0 = \operatorname{codim}(T_m\mathcal{S}_- + T_m\mathcal{U}_+)$. Therefore $\dim \operatorname{coker} D_{\gamma_0}\sigma_0 = 0$ and hence $\dim\ker D_{\gamma_0}\sigma_0 = 0$. Thus we conclude that σ_0 intersects the zero-section transversely.

It follows from Proposition 2.6 that there exists $\nu_0 > 0$ such that each solution of the equation $\sigma_\nu(\gamma_\nu) = 0,\ \nu \in (0, \nu_0)$ is contained in a C^0-neighbourhood $U(\gamma_0)$ of some $\gamma_0 \in \sigma_0^{-1}(0)$. Notice that the linearization of σ_ν at γ_ν can be written in the form

$$\begin{aligned} D_{\gamma_\nu}\sigma_\nu(\xi) &= \nabla_{\dot\gamma_\nu}\xi + \cos\theta_\nu \nabla_\xi v_0 + \sin\theta_\nu \nabla_\xi v_1 \\ &= \cos\theta_\nu\big(\nabla_\xi v_0 - \nabla_{v_0}\xi\big) + \sin\theta_\nu\big(\nabla_\xi v_1 - \nabla_{v_1}\xi\big). \end{aligned}$$

where $\xi \in W^{1,2}(\gamma_\nu^* TM)$. Hence, redenoting ν_0 if necessary, we can assume that the linearization of σ_ν is non-degenerate at each $\gamma_\nu \in \Gamma_\nu(m_-; m_+)$ contained in $\bigcup_{\gamma_0} U(\gamma_0)$ for $\nu \in (0, \nu_0)$, since $\#\sigma_0^{-1}(0) = \# L_- \cap L_+ < \infty$. Thus σ_ν intersects the zero-section transversely provided $\nu \le \nu_0$.

Consider σ_ν as a section of $\pi^*\mathcal{E}$, where $\pi\colon W^{1,2}_{m_-;\,m_+} \times \mathbb{R}_\nu \to \mathbb{R}_\nu$ is the canonical projection. Then σ_ν is continuous and satisfies the hypothesis of the implicit function theorem. Therefore, $\{(\gamma,\nu) \mid \sigma_\nu(\gamma)=0,\ \nu\in[0,\nu_0)\}$ is homeomorphic to $\sigma_0^{-1}(0)\times[0,\nu_0)$. This establishes the bijective correspondence between $\Gamma_0(m_-;m_+)$ and $\Gamma_\nu(m_-;m_+)$. □

Corollary B.3. *If L_+ and L_- intersect transversely in $f^{-1}(0)$, then there exists $\nu_0>0$ such that hypothesis (H2) holds provided $\nu\le\nu_0$.*

Acknowledgements

I thank S. Bauer, S. Donaldson, V. Pidstrygach, V. Rabinovich, Y. Ruan, D. Salamon and P. Seidel for helpful discussions and unknown referees for suggested improvements. My thanks also to B.Himpel for reading the draft of this paper. I acknowledge the financial support of the German Research Foundation (DFG) and the hospitality of Imperial College London, where part of this work was carried out.

References

[1] M. Atiyah, V. Patodi and I. Singer, *Spectral asymmetry and Riemannian geometry. III.* Math. Proc. Cambridge Philos. Soc., **79**(1):71–99, 1976.

[2] A. Banyaga and D. Hurtubise, *The Morse-Bott inequalities via a dynamical systems approach.* Ergodic Theory Dynam. Systems, **29**(6): 1693–1703, 2009.

[3] R. Beals, *Characterization of pseudodifferential operators and applications.* Duke Math. J., **44**(1):45–57, 1977.

[4] R. Bryant and S. Salamon, it On the construction of some complete metrics with exceptional holonomy. Duke Math. J., **58**(3):829–850, 1989.

[5] K. Cieliebak, A. Gaio and D. Salamon, *J-holomorphic curves, moment maps, and invariants of Hamiltonian group actions.* Internat. Math. Res. Notices, (16):831–882, 2000.

[6] K. Corlette, *Flat G-bundles with canonical metrics.* J. Differential Geom., **28**(3):361–382, 1988.

[7] S. Donaldson, *Twisted harmonic maps and the self-duality equations.* Proc. London Math. Soc. (3), **55**(1):127–131, 1987.

[8] S. Donaldson and E. Segal, *Gauge theory in higher dimensions, II.* In *Surveys in differential geometry. Volume XVI. Geometry of special holonomy and related topics*, volume 16 of *Surv. Differ. Geom.*, pages 1–41. Int. Press, Somerville, MA, 2011.

[9] H. Fan, T.J. Jarvis and Y. Ruan, *The Witten equation, mirror symmetry and quantum singularity theory.* `arXiv:0712.4021v3`.

[10] H. Fan, *Half de Rham complexes and line fields on odd-dimensional manifolds.* Trans. Amer. Math. Soc., **348**(8):2947–2982, 1996.

[11] A. Floer, *The unregularized gradient flow of the symplectic action.* Comm. Pure Appl. Math., **41**(6):775–813, 1988.

[12] D. Freed, *Remarks on Chern-Simons theory.* Bull. Amer. Math. Soc. (N.S.), **46**(2):221–254, 2009.

[13] A. Gaio and D. Salamon, *Gromov-Witten invariants of symplectic quotients and adiabatic limits.* J. Symplectic Geom., **3**(1):55–159, 2005.

[14] A. Haydys, *Gauge theory, calibrated geometry and harmonic spinors.* J. Lond. Math. Soc., II. Ser., **86**(2): 482–498, 2012.

[15] N. Hitchin, *The self-duality equations on a Riemann surface.* Proc. Lond. Math. Soc., III. Ser., **55**:59–126, 1987.

[16] A. Kapustin and E. Witten, *Electric-magnetic duality and the geometric Langlands program.* Commun. Number Theory Phys., **1**(1):1–236, 2007.

[17] D. McDuff and D. Salamon, *J-holomorphic curves and symplectic topology*, volume 52 of *American Mathematical Society Colloquium Publications.* American Mathematical Society, Providence, RI, 2004.

[18] V. Rabinovich and S. Roch, *Wiener algebras of operators, and applications to pseudodifferential operators.* Z. Anal. Anwendungen, **23**(3): 437–482, 2004.

[19] V. Rabinovich and S. Roch, *Agmon's type estimates of exponential behavior of solutions of systems of elliptic partial differential equations. Applications to Schrodinger, Moisil-Theodorescu and Dirac operators.* `arXiv:0802.3963`, 2008.

[20] J. Robbin and D. Salamon, *The Maslov index for paths.* Topology, **32** (4):827–844, 1993.

[21] J. Robbin and D. Salamon, *Asymptotic behaviour of holomorphic strips.* Ann. Inst. H. Poincaré Anal. Non Linéaire, **18**(5):573–612, 2001.

[22] D. Salamon, *Morse theory, the Conley index and Floer homology.* Bull. London Math. Soc., **22**(2):113–140, 1990.

[23] D. Salamon, *Lectures on Floer homology.* In *Symplectic geometry and topology (Park City, UT, 1997)*, volume 7 of *IAS/Park City Math. Ser.*, pages 143–229. Amer. Math. Soc., Providence, RI, 1999.

[24] P. Seidel, *Vanishing cycles and mutation.* In *European Congress of Mathematics, Vol. II (Barcelona, 2000)*, volume 202 of *Progr. Math.*, pages 65–85. Birkhäuser, Basel, 2001.

[25] P. Seidel, *Fukaya categories and Picard-Lefschetz theory.* Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2008.

[26] M. A. Shubin, *Spectral theory of elliptic operators on noncompact manifolds.* Astérisque, (207):5, 35–108, 1992. Méthodes semi-classiques, Vol. 1 (Nantes, 1991).

[27] C. Vafa and E. Witten, *A strong coupling test of S-duality.* Nuclear Phys. B, **431**(1-2):3–77, 1994.

[28] E. Witten, *Analytic continuation of Chern-Simons theory.* In *Chern-Simons gauge theory: 20 years after*, volume 50 of *AMS/IP Stud. Adv. Math.*, pages 347–446. Amer. Math. Soc., Providence, RI, 2011.

[29] E. Witten, *A new look at the path integral of quantum mechanics.* In *Surveys in differential geometry. Volume XV. Perspectives in mathematics and physics*, volume 15 of *Surv. Differ. Geom.*, pages 345–419. Int. Press, Somerville, MA, 2011.

[30] E. Witten, *Fivebranes and knots.* Quantum Topology, **3**(1):1–137, 2012.

Department of Mathematics, University Of Bielefeld
Postfach 100131, D-33501 Bielefeld, Germany
E-mail address: haydys@math.uni-bielefeld.de

Received October 9, 2012
Accepted August 23, 2013